

Université de Montréal

**Impacts des réarrangements génomiques chloroplastiques
sur l'apparition des phénotypes de variégation chez
*Arabidopsis thaliana***

par

Eric Zampini

Département de biochimie et de médecine moléculaire

Faculté de Médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Docteur en biochimie

Mai 2015

© Eric Zampini, 2015

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

**Impacts des réarrangements génomiques chloroplastiques
sur l'apparition des phénotypes de variégation chez
*Arabidopsis thaliana***

présentée par :

Eric Zampini

A été évaluée par un jury composé des personnes suivantes :

Dr. Pascal Chartrand, président-rapporteur

Dr. Normand Brisson, directeur de recherche

Dr. Pascale Beaugard, membre du jury

Dr. Hugo Germain, examinateur externe

Dr. Jacques Archambault, représentant du doyen

Résumé

Contrairement à la plupart des eucaryotes non-photosynthétiques, les végétaux doivent assurer la stabilité d'un génome additionnel contenu dans le plastide, un organite d'origine endosymbiotique. Malgré la taille modeste de ce génome et le faible nombre de gènes qu'il encode, celui-ci est absolument essentiel au processus de photosynthèse. Pourtant, même si ce génome est d'une importance cruciale pour le développement de la plante, les principales menaces à son intégrité, ainsi que les conséquences d'une déstabilisation généralisée de sa séquence d'ADN, demeurent largement inconnues.

Dans l'objectif d'élucider les conséquences de l'instabilité génomique chloroplastique, nous avons utilisé le mutant *why1why3pollb* d'*Arabidopsis thaliana*, qui présente d'importants niveaux de réarrangements génomiques chloroplastiques, ainsi que la ciprofloxacine, un composé induisant des brisures double-brins dans l'ADN des organites. Ceci nous a permis d'établir qu'une quantité importante de réarrangements génomiques provoque une déstabilisation de la chaîne de transport des électrons photosynthétique et un grave stress oxydatif associé au processus de photosynthèse. Étonnamment, chez *why1why3pollb*, ces hautes concentrations d'espèces oxygénées réactives ne mènent ni à la perte de fonction des chloroplastes affectés, ni à la mort cellulaire des tissus. Bien au contraire, ce déséquilibre rédox semble être à l'origine d'une reprogrammation génique nucléaire permettant de faire face à ce stress photosynthétique et conférant une tolérance aux stress oxydatifs subséquents.

Grâce à une nouvelle méthode d'analyse des données de séquençage de nouvelle génération, nous montrons également qu'un type particulier d'instabilité génomique, demeuré peu caractérisé jusqu'à maintenant, constitue une des principales menaces au maintien de l'intégrité génomique des organites, et ce, tant chez *Arabidopsis* que chez l'humain. Ce type d'instabilité génomique est dénommé réarrangement de type U-turn et est vraisemblablement associé au processus de réplication. Par une approche génétique, nous démontrons que les protéines chloroplastiques WHY1, WHY3 et RECA1 empêchent la

formation de ce type d'instabilité génomique, probablement en favorisant la stabilisation et le redémarrage des fourches de réplication bloquées. Une forte accumulation de réarrangements de type U-turn semble d'ailleurs être à l'origine d'un sévère trouble développemental chez le mutant *why1why3recA1*. Ceci soulève de nombreuses questions quant à l'implication de ce type d'instabilité génomique dans de nombreux troubles et pathologies possédant une composante mitochondriale.

Mots-Clés : biologie végétale, chloroplaste, mitochondrie, ROS, Whirly, RecA, instabilité génomique, réplication, recombinaison.

Abstract

In contrast to most non-photosynthetic eukaryotes, plants must ensure the stability of an additional genome contained within the plastid organelle. Despite the small size of the plastid genome and its low gene content, this genome is nevertheless absolutely essential for photosynthesis and plant energy metabolism. In spite of this, the main threats this genome encounters and their underlying consequences remain poorly understood.

To evaluate the consequences of generalized plastid genome instability, we use the *why1why3pollb Arabidopsis thaliana* mutant line, which exhibits elevated levels of plastid genome rearrangements, and ciprofloxacin, a compound that induces double strand-breaks within organelle DNA. We demonstrated that high levels of plastid genome rearrangements lead to a decrease in photosynthetic electron transport chain efficiency and to a severe photosynthesis-associated oxidative stress. Surprisingly, these high levels of reactive oxygen species are neither associated to a loss of chloroplast function, nor to cell death. Instead, this redox imbalance seems to initiate a nuclear genetic expression remodelling that allows adaptation to this photosynthetic stress and confers tolerance to subsequent oxidative stresses.

Using a novel approach for the analysis of next-generation sequencing data, we have also shown that a poorly characterized type of genomic instability constitutes one of the main threats to organelle genomic integrity, both in *Arabidopsis* and human. We demonstrate that this particular type of genomic instability, named U-turn-like DNA rearrangement, is most probably associated to errors during the replication process. Also, a genetic approach revealed that the chloroplast-localized proteins WHY1, WHY3 and RECA1 all act to repress this type of genomic instability, probably by stabilizing and stimulating the accurate restart of collapsed replication forks. A strong accumulation of U-turn-like rearrangements is notably associated to severe developmental defects in the *why1why3reca1* mutant line. This raises the question of whether this type of genomic instability could be involved in the appearance of several mitochondria-associated pathologies.

Keywords : plant biology, chloroplast, mitochondria, ROS, Whirly, RecA, genomic instability, replication, recombination.

Table des matières

Résumé.....	v
Abstract.....	vii
Table des matières.....	ix
Liste des tableaux	xiv
Liste des figures	xv
Liste des abréviations	xviii
Remerciements.....	xxii
CHAPITRE 1. Introduction.....	2
1.1 Les plastides : types et fonctions	3
1.2 Origine endosymbiotique du plastide et implications	6
1.2.1 Origine endosymbiotique du plastide	6
1.2.2 Pourquoi transférer le matériel génétique du plastide vers le noyau ?	8
1.3 Le génome du plastide : caractéristiques, structure et organisation	10
1.3.1 Structure et contenu génique.....	10
1.3.2 Le nucléoïde du plastide	12
1.4 Expression génique dans le plastide.....	14
1.4.1 La machinerie de transcription	14
1.4.2 La machinerie de traduction	15
1.5 La réplication de l'ADN du plastide	16
1.5.1 La machinerie de réplication	16
1.5.2 Les mécanismes de réplication	17
1.6 Les mécanismes assurant la stabilité du génome du plastide	20
1.6.1 Les mécanismes de prévention et de réparation des dommages oxydatifs	21
1.6.1.1 Les mécanismes prévenant la production de ROS par la photosynthèse	21
1.6.1.2 La détoxification des ROS au chloroplaste.....	27
1.6.1.3 La réparation des dommages oxydatifs de l'ADN chloroplastique.....	30
1.6.2 Les mécanismes de réparation des dommages à l'ADN induits par les rayons UV	32
1.6.3 Les mécanismes de réparation des brisures double-brin de l'ADN.....	34
1.6.3.1 La réparation des DSBs : les mécanismes de recombinaison homologue.....	35

1.6.3.2 La réparation des DSBs : les mécanismes de religation.....	38
1.6.3.3 La réparation des DSBs à une seule extrémité	40
1.6.3.4 Les protéines assurant la réparation des DSBs dans le chloroplaste	42
1.7 Objectifs et approche expérimentale.....	49
CHAPITRE 2. Plastid Genome Instability Leads to Reactive Oxygen Species Production and Plastid-to-Nucleus Retrograde Signaling in <i>Arabidopsis</i>	52
2.1 Mise en contexte	54
2.2 Abstract.....	55
2.3 Introduction	55
2.4 Results	58
2.4.1 Plastid Genome Instability Negatively Impacts Photosynthetic Efficiency.....	58
2.4.2 Plastid Genome Instability Leads to Generation of ROS in the Chloroplast.....	60
2.4.3 The <i>why1why3pollb-1</i> Phenotype Depends on Elevated ROS Levels.....	64
2.4.4 Yellow-Variegated Sectors in <i>why1why3pollb-1</i> Appear Photosynthetically Active and Are not Related to Cell Death.....	65
2.4.5 The <i>why1why3pollb-1</i> Yellow-Variegated Phenotype and Rearrangements Are not Maternally Inherited	67
2.4.6 <i>why1why3pollb-1</i> Exhibits a Photosynthetic Stress-Related Nuclear Genetic Reprogramming	68
2.4.7 The <i>why1why3pollb-1</i> Nuclear Genetic Reprogramming Closely Resembles ROS, Abiotic Stress and <i>msh1</i> Responses	71
2.4.8 The <i>why1why3pollb-1</i> Nuclear Genetic Reprogramming Is not Observed Under Low-Light Conditions	72
2.4.9 Genetic Reprogramming in <i>why1why3pollb-1</i> Allows Adaptation to High Light Stress.....	74
2.5 Discussion	76
2.5.1 Plastid Genome Instability Leads to Impairment of Photosynthesis and Increased Chloroplast ROS Production	76
2.5.2 ptDNA Rearrangements in <i>why1why3pollb-1</i> Do not Lead to “Permanently-Defective” Chloroplasts	77
2.5.3 The ROS-Associated <i>why1why3pollb-1</i> Nuclear Genetic Reprogramming Promotes Acquired Resistance to High Light.....	78
2.6 Materials and methods.....	80
2.6.1 Plant Material and Growth Conditions	80

2.6.2 Detection of ptDNA Rearrangements	81
2.6.3 Measurement of PSII Efficiency	81
2.6.4 Chlorophyll Extraction and Quantification	81
2.6.5 Chloroplast Ultrastructure Imaging	81
2.6.6 ROS Detection, Quantification and Microscopic Analyses	82
2.6.7 Anthocyanin Extraction and Quantification	82
2.6.8 RNA Isolation, RNA-Sequencing and Quantitative RT-PCR Analysis	82
2.6.9 Accession Numbers.....	83
2.7 Acknowledgements	84
2.8 Author contributions	84
2.9 Supplementary data	85
2.10 Supplemental materials and methods	95
2.10.1 Plant Material and Growth Conditions	95
2.10.2 Measurement of PSII Efficiency	95
2.10.3 Detection of ptDNA Rearrangements	95
2.10.4 Trypan blue staining	95
2.10.5 Quantitative RT-PCR Analysis	96
CHAPITRE 3. Caractérisation des relations génétiques entre les gènes <i>RECA1</i>, <i>POLIB</i>, <i>WHY1</i> et <i>WHY3</i>	98
3.1 Mise en contexte	99
3.2 Introduction	100
3.3 Résultats et discussion.....	102
3.3.1 Caractérisation de la relation génétique entre <i>RECA1</i> et <i>POLIB</i>	102
3.3.2 <i>RECA1</i> est un interacteur génétique des Whirly chloroplastiques	104
3.3.3 Les gènes <i>RECA1</i> , <i>WHY1</i> et <i>WHY3</i> travaillent de concert au maintien de la stabilité génomique chloroplastique	107
3.3.4 Les mutants <i>why1why3pollb</i> et <i>why1why3reca1</i> possèdent des patrons similaires de reprogrammation génique nucléaire.....	109
3.3.5 Le phénotype de variégation blanche de <i>why1why3reca1</i> n'est pas affecté par l'intensité lumineuse	110
3.4 Conclusion.....	112
3.5 Méthodes.....	113

CHAPITRE 4. Organelle DNA rearrangement mapping reveals U-turn-like inversions as a major source of genomic instability in <i>Arabidopsis</i> and humans.....	114
4.1 Mise en contexte	116
4.2 Abstract.....	117
4.3 Introduction	117
4.4 Results	119
4.4.1 Short-range rearrangements are abundant in organelle genomes.....	119
4.4.2 Microhomology and non-microhomology repair happen at similar rates in wild-type organelles.....	121
4.4.3 U-turn-like rearrangements are ubiquitous among organelle short-range rearrangements	124
4.4.4 Whirly proteins protect the plastid genome from MHMR but not from NHEJ	125
4.4.5 Whirly proteins, POLIB and RECA1 all act to maintain stability in the plastid genome.....	127
4.4.6 U-turn-like rearrangements are associated to replication stress.....	130
4.5 Discussion	132
4.6 Methods.....	136
4.6.1 Plant material and growth conditions	136
4.6.2 DNA isolation and DNA-seq	136
4.6.3 Publicly available Illumina whole-genome sequencing datasets	137
4.6.4 Enrichment for reads with potential junctions.....	137
4.6.5 Analysis of rearranged reads	138
4.6.6 Plastid sequencing coverage analysis	140
4.6.7 Pulse-field gel electrophoresis (PFGE) analysis of the plastid genome	140
4.6.8 Quantitative PCR analysis of ptDNA levels	141
4.6.9 Quantitative PCR analysis of RECA1 expression in <i>reca1</i> mutants	141
4.6.10 DATA ACCESS	142
4.7 Acknowledgements	142
4.8 Author contributions	142
4.9 Supplementary data	143
Chapitre 5. Discussion	154
5.1 L'instabilité génomique chloroplastique est associée à un déséquilibre photosynthétique .	155

5.2 Les protéines WHY1, WHY3, POLIB et RECA1 travaillent toutes au maintien de la stabilité génomique chloroplastique.....	158
5.3 Le processus de réplication menace l'intégrité génomique des chloroplastes et des mitochondries.....	161
5.3.1 : Les mécanismes de redémarrage des fourches de réplication bloquées.....	161
5.3.2 Évidences d'un mécanisme de redémarrage des fourches de réplication de type U-turn dans le chloroplaste et dans la mitochondrie.....	164
5.3.3 Rôles probables de RECA1 et des protéines Whirly dans le redémarrage des fourches de réplication	167
5.3.4 Les réarrangements de type U-turn semblent être une caractéristique intrinsèque de l'ADN des organites : les implications d'une telle structure	169
5.3.5 Les conséquences possibles d'une accumulation excessive de réarrangements de type U-turn dans l'ADN chloroplastique.....	173
5.3.6 Perspectives et expériences futures.....	175
Conclusion	179
Références.....	181

Liste des tableaux

Tableau I. Primer Pairs Used for the Specific Amplification of MMBIR-Associated ptDNA Rearrangements.	94
Tableau II. Primer Pairs Used for Quantitative RT-PCR experiments.....	94
Tableau III. Workflow Statistics for <i>Arabidopsis</i> plastid DNA rearrangements.....	152
Tableau IV. Workflow Statistics for <i>Arabidopsis</i> ecotypes Ts-1 and Ws-2 plastid DNA rearrangements.	152
Tableau V. Workflow Statistics for <i>Arabidopsis</i> mitochondria DNA rearrangements.	152
Tableau VI. Workflow Statistics for human brain mitochondria DNA rearrangements.....	153
Tableau VII. Workflow Statistics for human liver mitochondria DNA rearrangements.....	153
Tableau VIII. Workflow Statistics for <i>E. coli</i> DNA rearrangements.	153

Liste des figures

Figure 1. Représentation schématique d'un chloroplaste.	3
Figure 2. Représentation schématique des types de plastides et de leur différenciation.	4
Figure 3. Représentation de la contribution cyanobactérienne au génome d' <i>Arabidopsis</i>	7
Figure 4. Représentation circulaire du génome du plastide d' <i>Arabidopsis</i> et classification fonctionnelle de chacun des gènes.	12
Figure 5. Représentation schématique des modèles de la boucle D double et du cercle roulant.....	19
Figure 6. Les voies de signalisation rétrograde chloroplaste-noyau et quelques-uns de leurs principaux effets sur la reprogrammation génétique nucléaire.	25
Figure 7. Les mécanismes et enzymes participant à la détoxification des ROS dans le chloroplaste.	29
Figure 8. L'oxydation des bases azotées génère de nombreux dérivés de bases azotées, dont la 8-oxoguanine et la thymine glycol.	31
Figure 9. Les rayons UV-B mènent à la formation de liens covalents entre des pyrimidines adjacentes.	33
Figure 10. La réparation d'un DSB par recombinaison homologue.	36
Figure 11. La réparation d'un DSB par religation des deux extrémités du bris.....	39
Figure 12. La réparation d'un DSB à une seule extrémité par les processus de BIR et de MMBIR.....	41
Figure 13. Les protéines Whirly favorisent la réparation adéquate des DSBs en stabilisant l'ADNss et en limitant les appariements non-spécifiques.	46
Figure 14. Plastid DNA Rearrangements in <i>why1why3pollb-1</i> Lead to Photosynthetic Impairment and Yellow Variegation.	59
Figure 15. <i>why1why3pollb-1</i> Produces Elevated Levels of Chloroplast ROS in a Light-Dependent Manner.	61
Figure 16. Ciprofloxacin-Treated Seedlings Show Signs of Bleaching Associated With Increased Chloroplast ROS Production.	63
Figure 17. Growth under Low Light Conditions Rescues the <i>why1why3pollb-1</i> Phenotype.	65
Figure 18. The <i>why1why3pollb-1</i> Yellow-Variegated Sectors Are not Devoid of Photosynthetically Active Chloroplasts.	66
Figure 19. The <i>why1why3pollb-1</i> ptDNA Rearrangements and Phenotype Are not Maternally Inherited.	67
Figure 20. The <i>why1why3pollb-1</i> Mutant Exhibits a Complex Nuclear Genetic Reprogramming.	70

Figure 21. The <i>why1why3pollb-1</i> Nuclear Genetic Reprogramming Is Not Observed under Low-Light Conditions.	73
Figure 22. The Early Oxidative Stress in <i>why1why3pollb-1</i> Leads to Acquired Resistance to High Light.	75
Figure 23. Microhomology-Mediated Break-Induced Replication (MMBIR) Leads to DNA Rearrangements.	85
Figure 24. The <i>why1why3pollb-1</i> Mutant Develops Distinct Yellow and Green Sectors on Its Leaves.	86
Figure 25. Some <i>why1why3pollb-1</i> Plants Exhibit Yellow and White Variegation Simultaneously.	87
Figure 26. Ciprofloxacin Negatively Impacts Photosynthetic Efficiency in <i>Arabidopsis</i>	88
Figure 27. <i>why1why3</i> and <i>why1why3pollb-1</i> Plants Are Hypersensitive to Paraquat.	89
Figure 28. The Abundance of MMBIR-Associated ptDNA Rearrangements Is Little Affected by the Light-Regimen in <i>why1why3pollb-1</i>	90
Figure 29. Paraquat and Atrazine do not induce MMBIR-Associated ptDNA Rearrangements.....	91
Figure 30. Elevated ROS Levels Do not Lead to Cell Death in <i>why1why3pollb-1</i>	92
Figure 31. The <i>why1why3pollb-1</i> Nuclear Genetic Reprogramming Is Not Due to the <i>pollb-1</i> and <i>why1why3</i> Genetic Backgrounds.....	93
Figure 32. La double mutation <i>reca1pollb</i> provoque une grave sensibilité à la CIP, mais ne provoque pas d'accumulation de réarrangements génomiques.	103
Figure 33. Le gène <i>RECA1</i> est un interacteur génétique des Whirly chloroplastiques.	106
Figure 34. Les mutations <i>why1why3reca1</i> et <i>why1why3pollb</i> mènent à une forte accumulation de réarrangements génomiques chloroplastiques et à une signalisation rétrograde du chloroplaste au noyau.	108
Figure 35. Le phénotype de variégation blanche n'est pas affecté par l'intensité lumineuse chez <i>why1why3reca1</i>	111
Figure 36. Global portrait of organelle genome rearrangements in <i>Arabidopsis</i> and human.	120
Figure 37. Analysis of organelle genome rearrangements in <i>Arabidopsis</i> and human.....	123
Figure 38. Global portrait of plastid genome rearrangements in <i>Arabidopsis</i> mutant lines <i>pollb</i> , <i>reca1</i> and <i>why1why3</i>	126
Figure 39. Global portrait of plastid genome rearrangements in <i>Arabidopsis</i> mutant lines <i>reca1pollb</i> , <i>why1why3pollb</i> and <i>why1why3reca1</i>	129
Figure 40. Plastid DNA sequencing coverage curves for <i>Arabidopsis</i> lines WT, <i>why1why3pollb</i> and <i>why1why3reca1</i>	131
Figure 41. Model of microhomology-mediated U-turn-like inversions.....	134

Figure 42. Level of plastid DNA rearrangements in three <i>Arabidopsis</i> ecotypes.....	143
Figure 43. Characterization of the <i>reca1-1</i> and <i>reca1-2</i> T-DNA insertion mutant lines.....	144
Figure 44. Characterization of the <i>why1why3reca1</i> phenotype.	145
Figure 45. Plastid DNA sequencing coverage curves for WT, <i>pollb</i> , <i>reca1</i> , <i>reca1pollb</i> and <i>why1why3</i> <i>Arabidopsis</i> mutant plants.....	146
Figure 46. Plastid DNA quantification at three locations of the genome in WT, <i>why1why3pollb</i> and <i>why1why3reca1</i> plants.	147
Figure 47. Visualization of the distribution of the various forms of the plastid genome.	148
Figure 48. Schematic representation of the analysis workflow.	149
Figure 49. Comparison of the techniques used to detect genome rearrangements.	150
Figure 50. Local mapping of read pairs associated to specific genome rearrangements.	151
Figure 51. L'instabilité génomique chloroplastique est associée à un déséquilibre rédox et à une reprogrammation génique nucléaire.....	157
Figure 52. Les protéines Whirly, POLIB et RECA1 favorisent la stabilité génomique chloroplastique.	159
Figure 53. La régression des fourches de réplication bloquées et les mécanismes de redémarrage de la réplication.	162
Figure 54. Distribution des microhomologies peu espacées dans le génome chloroplastique d' <i>Arabidopsis</i>	165
Figure 55. Conséquences possibles de la ligation du brin continu au brin discontinu lors d'un réarrangement de type U-turn.	170
Figure 56. L'ADN mitochondrial du cœur humain est organisé en structures hautement branchées.....	172

Liste des abréviations

ADN :	Acide désoxyribonucléique
ADNc :	ADN complémentaire
ADNpt :	ADN du plastide
ADNss :	ADN simple-brin
ARN :	Acide ribonucléique
ATP :	Adénosine triphosphate
BIR :	Break-Induced Replication
bp :	Paires de bases
CPD :	Cyclobutane pyrimidine dimer
DSB :	Brisures double-brins de l'hélice d'ADN
DSBR :	Double-Strand Break Repair
IRs :	Grandes séquences inversées et répétées du plastide
kb :	Kilobase
LHC :	Light harvesting complexes
MEcPP :	Méthylérythritol cyclodiphosphate
MHMR :	Microhomology-mediated recombination
MMBIR :	Microhomology-Mediated Break-induced Replication
MMEJ :	Microhomology-Mediated End Joining
NEP :	Nuclear-encoded RNA polymerase
NHEJ :	Non-Homologous End Joining
PAP :	3'-phosphoadenosine 5'-phosphate
PCR :	Polymerase Chain Reaction

PEP :	Plastid-encoded RNA Polymerase
PET :	Chaîne de transport des électrons photosynthétique
PFGE :	Électrophorèse en champ pulsé
PSI, PSII :	Photosystèmes I et II
ROS :	Espèces oxygénées réactives
SDSA :	Synthesis-Dependent Strand Annealing
Site AP :	Site apurinique/apyrimidique
SSA :	Single-Strand Annealing

It seems to me that the natural world is the greatest source of excitement; the greatest source of visual beauty; the greatest source of intellectual interest. It is the greatest source of so much in life that makes life worth living.

- David Attenborough

Remerciements

Avant tout, je tiens à remercier mon directeur de recherche Dr. Normand Brisson. Merci Normand de m'avoir accueilli dans votre laboratoire le temps d'un été pour un stage au baccalauréat. La belle équipe, le sujet de recherche et l'atmosphère motivante m'ont donné envie de poursuivre dans votre labo. Je vous remercie aussi d'avoir toujours été disponible et prêt à m'aider lorsque nécessaire. J'ai tellement appris de choses ces dernières années, ce fut toute une expérience pour moi ce passage au labo! Et aussi, merci pour toutes ces belles discussions sur tellement de sujets divers et ces sorties en voilier, cela a certainement contribué à façonner la belle équipe que nous avons formée!

Évidemment, des remerciements immenses vont à Étienne et Samuel. Je ne sais comment exprimer simplement ce qu'à représenter pour moi ces dernières années à vos côtés. Étienne, merci pour le camarade fidèle que tu as été, pour ton leadership, ta motivation et toutes les choses que tu m'as apprises. J'ai énormément apprécié que tu sois toujours prêt à travailler en équipe et à avoir des projets communs, c'était décidément une formule gagnante. Avec ton énergie débordante et ta bonne humeur contagieuse, c'était toujours un plaisir de travailler avec toi. À Sam, merci pour tous ces beaux moments passés ensemble. Quand j'ai travaillé avec toi le temps d'un stage d'été, je ne me doutais pas que je développerais une aussi belle amitié avec toi. Bonne chance avec la suite des projets de séquençage! Ils sont entre bonnes mains, je suis certain que tu vas nous impressionner! Je tiens aussi à vous remercier tous les deux pour votre contribution déterminante à ce dernier projet. On a fait face à beaucoup d'obstacles avec toute cette bio-info, mais ça a bien montré l'importance du travail d'équipe... Je garderai tellement de bons souvenirs des moments passés avec vous les gars, avec les congrès et tout le reste... Je vous souhaite la meilleure des chances en sciences et j'espère très fortement qu'on restera en contact encore longtemps! Merci pour ces belles années!

Merci à toi Jean-Sébastien pour m'avoir formé et pour ta patience malgré les nombreuses gaffes que j'ai pu faire à mon arrivée au laboratoire. Merci pour tous tes conseils et pour tout le temps que tu as pris pour m'aider, je l'ai beaucoup apprécié! Merci à toi Laurent pour tout le temps que tu as pris pour m'aider, surtout avec le FPLC... Une chose est certaine, tu m'as appris comment travailler efficacement! Alexandre, même si on ne s'est pas côtoyé très longtemps, je garderai de toi un excellent souvenir. Merci pour tes conseils et les discussions que l'on a eues pendant mon stage. Sébastien, merci pour les moments agréables qu'on a passés ensemble au laboratoire et pour l'aide que tu m'as apportée. Je suis content pour toi, tu sembles très heureux avec ton nouvel emploi. Je te souhaite la meilleure des chances pour

la suite! Merci à toi aussi Karine, pour le bel été que l'on a passé ensemble au laboratoire! Je suis bien content que tu fasses maintenant partie de mon cercle d'amis!

Je tiens également à remercier les Drs. David Morse et Muriel Aubry pour tous vos conseils et votre grand intérêt pendant chacun de mes comités de thèse. David, merci de m'avoir accueilli dans votre laboratoire lors de ce stage au baccalauréat, ce fut une très belle expérience. Merci également au Dr. Luc DesGroseillers pour ce stage en première année de baccalauréat. J'ai beaucoup aimé mon été et énormément appris lors de ce stage.

Un gros merci également à mes inséparables amis du baccalauréat en biochimie, Antoine, Benjamin et Simon. Cela fait maintenant huit ans que l'on se connaît et je suis certain que cette amitié durera encore bien longtemps! Merci pour tous ces beaux moments passés ensemble, il y a en aura encore bien d'autres, j'en suis sûr!

Merci aussi à tous mes amis au département, Mathieu, Emmanuelle, Lian, Eric, Thomas et tous les autres. Vous avez tous contribué à rendre mon passage au département des plus agréables! Je garderai d'excellents souvenirs de la fin de semaine de kayak, du ski, des randonnées et de toutes les autres activités que nous avons faites ensemble. Bonne continuation à vous tous et à tous ceux que j'ai pu oublier! Merci aussi à tous les employés au département pour votre aide. Merci aussi à Sougata, Steve et Mathieu pour le bon temps passé ensemble au laboratoire de David!

Merci également à ma famille et à belle-famille pour vos encouragements indéfectibles. Je sais que vous êtes tous très fiers même si vous ne comprenez pas toujours ce que je fais vraiment. Finalement, un merci immense à Marie-Christine. Merci de partager ma vie depuis bientôt trois ans. Merci pour ton soutien, tes encouragements et ton surprenant intérêt pour les chloroplastes! Tu m'inspires beaucoup. Je ne sais trop comment décrire tout ce que tu représentes pour moi. Merci pour tout, tout simplement...

CHAPITRE 1. Introduction

1.1 Les plastides : types et fonctions

Contrairement à la plupart des autres organismes vivants, les végétaux possèdent des plastides, une famille d'organites qui leur est propre et qui effectue une multitude de fonctions au sein de la cellule. La plus connue de ces fonctions est sans nul doute la photosynthèse, le processus bioénergétique à la base de la vie sur Terre par lequel les organismes photosynthétiques convertissent l'énergie solaire en énergie chimique. Par ce processus, le CO₂ atmosphérique est réduit en glucides selon l'équation générale suivante :



Le chloroplaste, le siège de la photosynthèse, est un plastide spécialisé constitué de trois membranes : la membrane externe, la membrane interne et les thylakoïdes. Alors que les membranes externe et interne définissent l'espace intermembranaire, la membrane interne entoure le stroma (Fig. 1). Quant aux thylakoïdes, ceux-ci consistent en un réseau de disques membranaires interconnectés retrouvé dans le stroma.

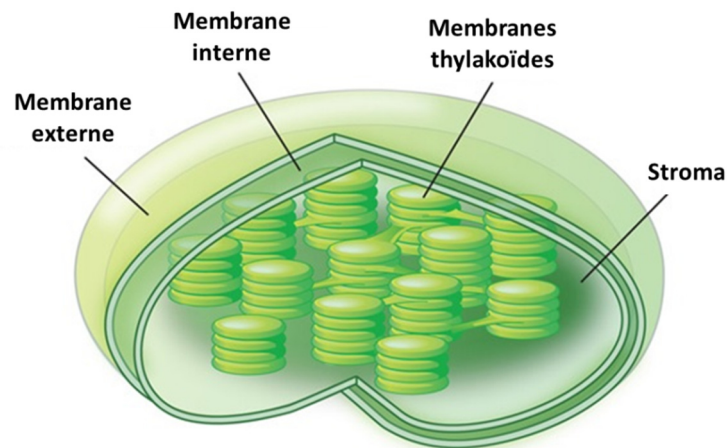


Figure 1. Représentation schématique d'un chloroplaste. Cette figure est adaptée de [1].

Le processus de photosynthèse débute au sein de la membrane des thylakoïdes, où les complexes d'antennes, généralement désignés « light harvesting complexes » (LHC), capturent des photons par le biais de pigments nommés chlorophylles. Cette excitation de la

chlorophylle provoque le transfert d'un électron vers des centres réactionnels photochimiques de la chaîne de transport des électrons photosynthétique (PET), nommés photosystèmes I et II (PSI et PSII). Cet électron provenant de la chlorophylle est subséquentement remplacé par un électron peu énergétique provenant d'une molécule d'eau. De manière similaire à la phosphorylation oxydative mitochondriale, le passage d'électrons au travers de la PET génère un gradient de protons de part et d'autre de la membrane des thylakoïdes, ce qui permet la production d'ATP. D'autre part, le fort pouvoir réducteur des électrons est également mis à profit pour générer du NADPH. L'ATP et le NADPH générés servent ensuite d'énergie et de pouvoir réducteur pour la transformation du CO_2 en sucres grâce, notamment, à l'enzyme ribulose-1,5-biphosphate carboxylase oxygénase (RuBisCO).

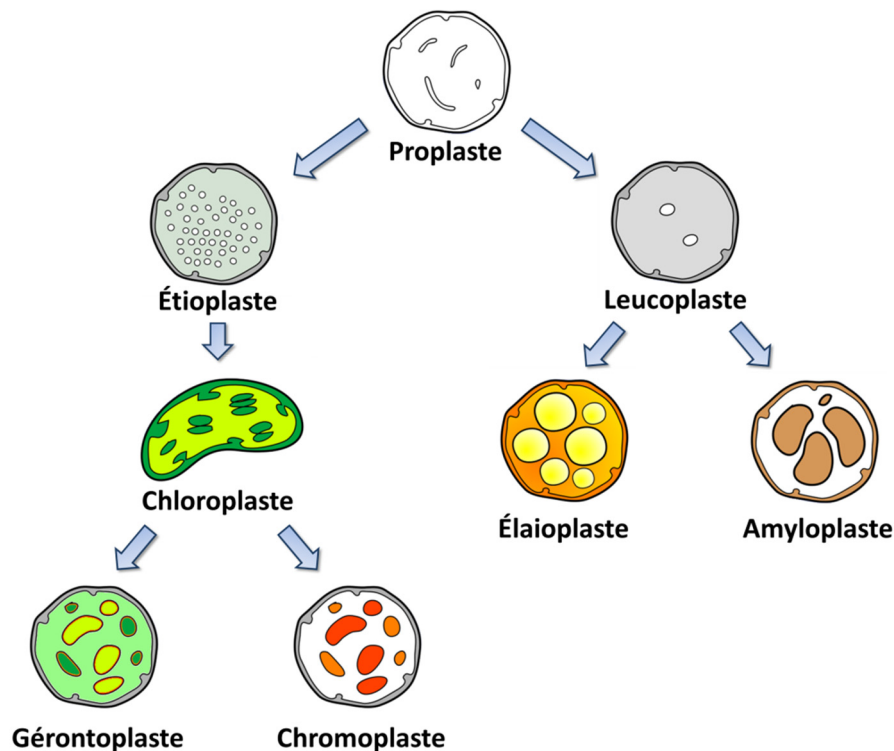


Figure 2. Représentation schématique des types de plastides et de leur différenciation.

Toutefois, les plastides ne font pas que participer au processus de photosynthèse. Bien au contraire, ceux-ci accomplissent une multitude d'autres tâches et sont impliqués dans de très

nombreuses voies de biosynthèse. En effet, ces organites participent à l'absorption de l'azote et du soufre en plus de contribuer à la synthèse des acides aminés, des lipides, de l'amidon et de la chlorophylle [2,3]. Par conséquent, les plastides se subdivisent en plusieurs types, chacun accomplissant certains rôles particuliers. Les proplastides, ou plastides non-différenciés, sont retrouvés dans les méristèmes, les tissus végétaux hautement réplicatifs, ainsi que dans les tissus reproductifs. Par une succession de différenciations, les proplastides sont à l'origine de tous les autres types de plastides (Fig. 2) et assurent la transmission de cette famille d'organites dans les cellules filles et les gamètes [revu dans 2]. Les proplastides génèrent ainsi les étioplastes, les précurseurs des chloroplastes, jusqu'à ce que la plante soit exposée à une source lumineuse. À ce moment, les étioplastes développent très rapidement un réseau de membranes thylakoïdes et assemblent la machinerie photosynthétique pour former des chloroplastes fonctionnels. Bien que les chloroplastes assurent de nombreuses fonctions précises au sein de la cellule, ceux-ci peuvent se redifférencier en deux autres types de plastides : les chromoplastes et les gérontoplastes [2]. Les chromoplastes, qui sont principalement retrouvés dans les fruits et les pétales, ont pour principale fonction le stockage de pigments colorés et servent, entre autres, d'attractant pour les pollinisateurs. Les gérontoplastes, quant à eux, sont retrouvés dans les tissus sénescents et servent au démantèlement sécuritaire de la machinerie photosynthétique afin de prévenir un stress oxydatif cellulaire (voir section 1.6.1.1). Par une voie de différenciation distincte, les proplastides engendrent aussi les leucoplastes, les précurseurs des élaïoplastes et des amyloplastides, des plastides spécialisés, respectivement, dans l'entreposage des huiles et de l'amidon [2]. Grâce à leur haute densité, les amyloplastides jouent aussi un rôle dans la perception de la gravité par la plante afin de promouvoir une croissance polarisée des tissus.

Les plastides étant les descendants d'une cyanobactérie ancestrale (voir section 1.2.1), l'étude de cette famille d'organites nécessite de garder en perspective leur origine procaryote et les caractéristiques qui y sont associées. Cette origine bactérienne se manifeste notamment par la structure et les caractéristiques du génome du plastide (voir

section 1.3), de même que par la machinerie et les mécanismes de transcription, de réplication et de réparation de l'ADN du plastide (ADNpt) (voir sections 1.4 à 1.6).

1.2 Origine endosymbiotique du plastide et implications

1.2.1 Origine endosymbiotique du plastide

La similarité existant entre les chloroplastes et les cyanobactéries a été initialement remarquée il y a plus d'un siècle [traduction anglaise dans 4]. Que ce soit par la coloration associée à la chlorophylle, la présence de thylakoïdes ou la capacité photosynthétique de chacun, ces trois particularités suggèrent un ancêtre commun. Il est maintenant généralement accepté qu'un phénomène d'endosymbiose, survenu il y a environ 1,5 milliards d'années, est à l'origine des plastides [revu dans 5,6]. Selon ce modèle, l'ancêtre des eucaryotes modernes aurait établi une relation endosymbiotique avec une α -protéobactérie, menant à l'émergence des mitochondries [revu dans 7]. Un descendant de cette lignée aurait ensuite phagocyté une cyanobactérie et acquérant, du même coup, la capacité d'effectuer la photosynthèse. Puisque les eucaryotes possèdent tous des mitochondries, il semble donc logique que ces deux phénomènes d'endosymbiose se soient déroulés dans cet ordre.

Plusieurs évidences soutiennent l'hypothèse que les plastides possèdent un ancêtre cyanobactérien. À ce sujet, l'analyse phylogénétique des ARNs ribosomiques 16S et 23S révèle que les plastides des plantes terrestres et des algues descendent d'une cyanobactérie ancestrale [8-10]. Cette parenté se perçoit également au niveau des gènes encodés dans le génome du plastide (voir section 1.3.1), plusieurs d'entre eux étant des homologues de gènes cyanobactériens. Il existe par ailleurs une forte similarité entre plusieurs protéines des plastides et des cyanobactéries [revu dans 5,10].

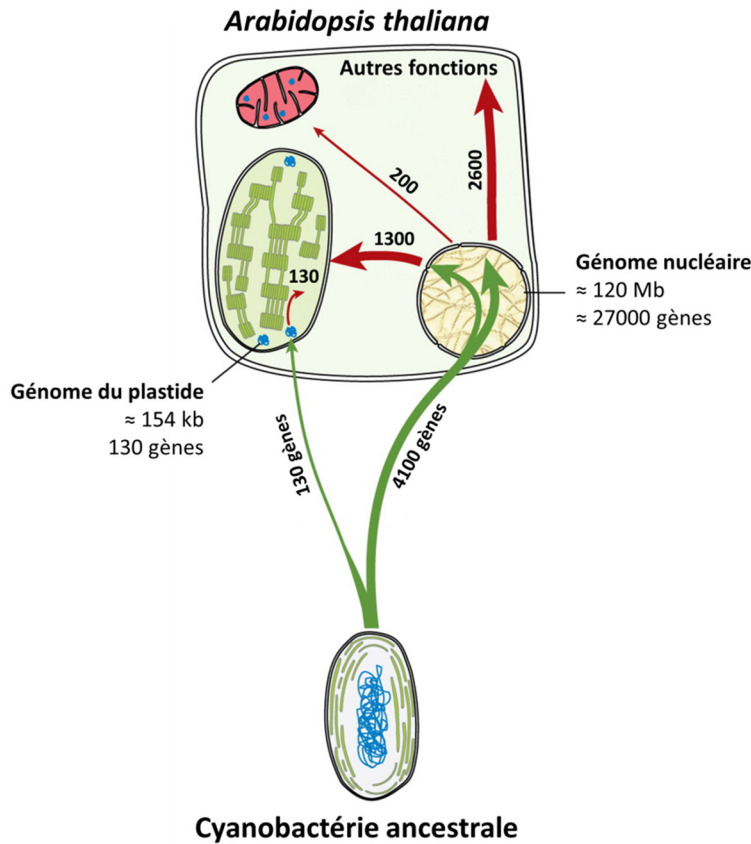


Figure 3. Représentation de la contribution cyanobactérienne au génome d'*Arabidopsis*. Les flèches vertes représentent l'ampleur de la contribution cyanobactérienne au génome nucléaire et du plastide. Les flèches rouges représentent la localisation cellulaire des produits des gènes d'origine cyanobactérienne. Cette figure est adaptée de [11].

Suite à cette endosymbiose, la coexistence au sein d'une même cellule du génome de l'hôte et du génome cyanobactérien a vraisemblablement rendu superflu le maintien de la totalité de ces gènes. Devenus peu utiles, de nombreux gènes du génome cyanobactérien ont ainsi disparu au fil de l'évolution [12]. D'autres ont été transférés au noyau lorsque les membranes du plastide étaient rompues et que des fragments d'ADNpt étaient intégrés dans le génome nucléaire [5,6,13]. Il est estimé que plus de 4000 gènes nucléaires sont d'origine cyanobactérienne chez *Arabidopsis* et, de ceux-ci, environ la moitié ont acquis des fonctions cellulaires ailleurs qu'au plastide [revu dans 11] (Fig. 3). Ce phénomène peut d'ailleurs être observé en laboratoire par l'insertion, dans le génome chloroplastique, d'un gène de

résistance ne fonctionnant que dans le contexte du génome nucléaire [14,15]. Ces expériences ont révélé que la transposition de fragments d'ADN_{pt} dans le noyau est particulièrement fréquente, ce taux atteignant un événement par 16 000 grains de pollen [14]. Suite à leur intégration dans le noyau, plusieurs de ces gènes ont acquis des promoteurs et des séquences permettant le ciblage des polypeptides qu'ils encodent au plastide [12]. Devenus superflus, les gènes de l'ancêtre cyanobactérien ont donc presque entièrement disparu du génome du plastide chez les eucaryotes modernes (voir section 1.3.1). Pour cette raison, la plupart des ≈ 2500 protéines localisées au plastide sont encodées dans le noyau et importées dans cet organe suite à leur synthèse dans le cytosol [revu dans 3,16].

1.2.2 Pourquoi transférer le matériel génétique du plastide vers le noyau ?

À priori, il apparaît raisonnable de croire que ce transfert massif de gènes vers le noyau aurait permis une régulation génique plus fine et plus complexe. Cette hypothèse semble néanmoins peu probable, puisqu'il existe un système élaboré de régulation génique au sein du plastide [revu dans 17] (voir section 1.4). En effet, la régulation transcriptionnelle des gènes encodés au plastide se trouve sous le contrôle de deux ARN polymérases distinctes, la « Nuclear-encoded RNA polymerase » (NEP) et la « Plastid-encoded RNA polymerase » (PEP). Puisque cette seconde polymérase est régulée par six facteurs de transcription sigmas de type procaryote encodés au noyau, l'expression des gènes du plastide se retrouve donc partiellement sous contrôle nucléaire [18]. D'autre part, une complexe régulation traductionnelle contrôle l'expression génique du plastide et ajoute un niveau supplémentaire de régulation. Cette relocalisation nucléaire de segments de l'ADN_{pt} ne semble donc pas être un préalable au contrôle précis et intégré de l'expression génique du plastide.

Puisqu'il semble peu probable que ce transfert de gènes du plastide vers le noyau se soit strictement produit pour des raisons de régulation génique, certaines hypothèses alternatives ont été suggérées. En raison de son rôle dans le processus de photosynthèse, le chloroplaste peut être sujet à des variations de son état rédox et, par le fait même, à d'éventuels dommages oxydatifs de ses constituants (voir section 1.6.1). Le transfert d'une

partie de son matériel génétique au noyau a, par conséquent, probablement permis d'éviter les mutations associées à de tels dommages [revu dans 6,17].

Il a aussi été suggéré que ce transfert aurait permis d'éviter une détérioration du génome du plastide par le processus de « Muller's Ratchet » [19]. Selon cette théorie, un génome se transmettant de manière asexuée accumule inévitablement des mutations délétères, celles-ci ne pouvant être épurées par recombinaison lors de la méiose. Puisque le plastide présente une transmission maternelle, le transfert de gènes au noyau aurait permis le passage d'une transmission de type clonale à une de type sexuée, prévenant une dégradation du matériel génétique [6,17]. Cette hypothèse a toutefois été fortement remise en question, notamment parce que le génome du plastide présente un taux de mutations plus faible que celui du génome nucléaire [20]. Il semble donc que certains mécanismes préviennent une telle dégradation du matériel génétique. La transmission maternelle des plastides n'est d'ailleurs pas absolue, une transmission paternelle étant parfois observée et occasionnant un certain niveau de recombinaison sexuée [revu dans 21]. De plus, ce faible taux de mutations de l'ADNpt suggère aussi que des mécanismes de stabilité génomique et de réparation de l'ADN sont à l'œuvre dans ce compartiment (voir section 1.6). À ce sujet, il a été démontré que des événements de conversion génique, un type particulier de recombinaison, préviennent l'accumulation de mutations ponctuelles [22].

Quoique ces hypothèses apportent quelques pistes de solution, les avantages que confère ce transfert de matériel génétique du plastide vers le noyau restent toutefois spéculatifs et peu compris. De plus, cet échange de matériel génétique soulève une seconde question peut-être aussi complexe : pour quelle raison ce transfert de gènes n'est-il pas allé jusqu'au bout, en relocalisant la totalité des gènes du plastide? Et pour quelle raison le plastide contient-il encore un génome?

Tel que souligné par Bock et Timmis, une très forte sélection explique probablement le maintien du génome du plastide [12]. Ceux-ci font d'ailleurs remarquer que la plupart des composantes de la machinerie photosynthétique encodée au plastide constituent le « cœur »

des centres réactionnels photochimiques, au contraire des composantes périphériques qui sont principalement encodées au noyau. Les composantes de « cœur » étant encodées au plastide et dirigées vers les membranes thylakoïdes, les centres réactionnels photochimiques ne peuvent donc pas s'assembler au sein d'autres membranes de la cellule et causer un stress oxydatif. D'autre part, la proximité du génome du plastide et de la PET permet une adaptation rapide de la capacité photosynthétique lors de fluctuations environnementales et contribue ainsi à prévenir un potentiel stress oxydatif [12].

1.3 Le génome du plastide : caractéristiques, structure et organisation

Tel que décrit à la section précédente, le génome du plastide est issu d'un ancêtre cyanobactérien et a perdu, ou transféré, la majorité de ses gènes vers le noyau au fil de l'évolution. Il en résulte un génome aux caractéristiques uniques et particulières, bien distinctes de la majorité des autres génomes. Le génome mitochondrial fait toutefois exception, celui-ci étant également issu d'un phénomène d'endosymbiose et partageant certaines caractéristiques avec le génome du plastide. La section suivante dresse un portrait des principales caractéristiques du génome du plastide retrouvé chez les plantes terrestres et les algues vertes. Ces organismes constituent la lignée « verte » du règne *Plantae* et incluent l'organisme modèle *Arabidopsis thaliana* qui est étudié dans ce travail.

1.3.1 Structure et contenu génique

Chez cette lignée « verte », le génome du plastide possède une très petite taille, celle-ci étant généralement comprise entre 120 et 160 kilobases (kb) et mesurant 154 kb chez *Arabidopsis thaliana* [revu dans 23]. Certaines algues vertes de la division des chlorophytes font toutefois exception, leur génome pouvant atteindre plusieurs centaines de kb en raison d'une abondance extrême de séquences répétées dans les régions intergéniques [24]. Le génome du plastide est habituellement représenté sous forme d'une molécule d'ADN circulaire comportant deux grandes séquences inversées et répétées (IRs) de longueur variable selon les espèces (≈ 26 kb chez *Arabidopsis thaliana*, Fig. 4). Alors que le nombre de gènes encodés

dans le génome entier est relativement constant entre les espèces, le nombre de gènes contenu dans les IRs est, quant à lui, variable et en grande partie responsable de la variabilité de la taille du génome entre espèces [revu dans 23]. Ces IRs définissent deux régions à simple copie de taille inégale : une petite (SSC, *small single-copy*) et une grande (LSC, *large single-copy*). Bien que certaines évidences démontrent que les IRs serviraient de matrice à des événements de recombinaison initiant la réplication (voir section 1.5.2), leur pertinence biologique reste néanmoins peu comprise. Il est toutefois intéressant de constater que le taux de mutations ponctuelles est plus faible au sein des IRs que dans les régions à simple copie [25], des mécanismes efficaces de conversion génique limitant probablement l'accumulation de mutations dans les IRs. Les IRs ne sont cependant pas les seules séquences répétées retrouvées dans le génome du plastide. De nombreuses séquences répétées très courtes (< 50 bp) sont aussi réparties sur l'ensemble du génome, et ce tout particulièrement chez les chlorophytes [24]. Ces courtes répétitions seraient responsables de l'accumulation de réarrangements génomiques [26-28], notamment lors de stress génotoxiques [29].

Tel que mentionné précédemment, peu de gènes d'origine cyanobactérienne subsistent dans le génome du plastide moderne. Chez *Arabidopsis*, le génome du plastide compte 130 gènes encodant presque exclusivement des composantes de la PET ou de la machinerie nécessaire à leur expression [revu dans 3]. Parmi ceux-ci, 34 encodent la majorité des polypeptides qui composent les quatre principaux complexes de la PET, soit les photosystèmes I et II, le cytochrome b_6f et l'ATP synthase [revu dans 3,23]. En plus de ces gènes, le génome du plastide d'*Arabidopsis* encode quatre ARNs ribosomiques, toujours retrouvés dans les IRs et donc dédoublés, 37 ARNs de transfert, quatre sous-unités de la « Plastid-encoded RNA polymerase » et 26 protéines ribosomales [3]. À l'exception de la grande sous-unité de Rubisco, *RBCL*, la totalité des gènes nécessaires à la fixation du carbone a été transférée au noyau. Quoique la distribution de ces gènes dans le génome varie entre les organismes, plusieurs ont tendance à rester groupés afin, notamment, de réunir des unités transcriptionnelles en opéron et de générer des polycistrons [3]. Des séquences Shine-Dalgarno initient d'ailleurs la traduction de plusieurs ARNs messagers du plastide [30,31].

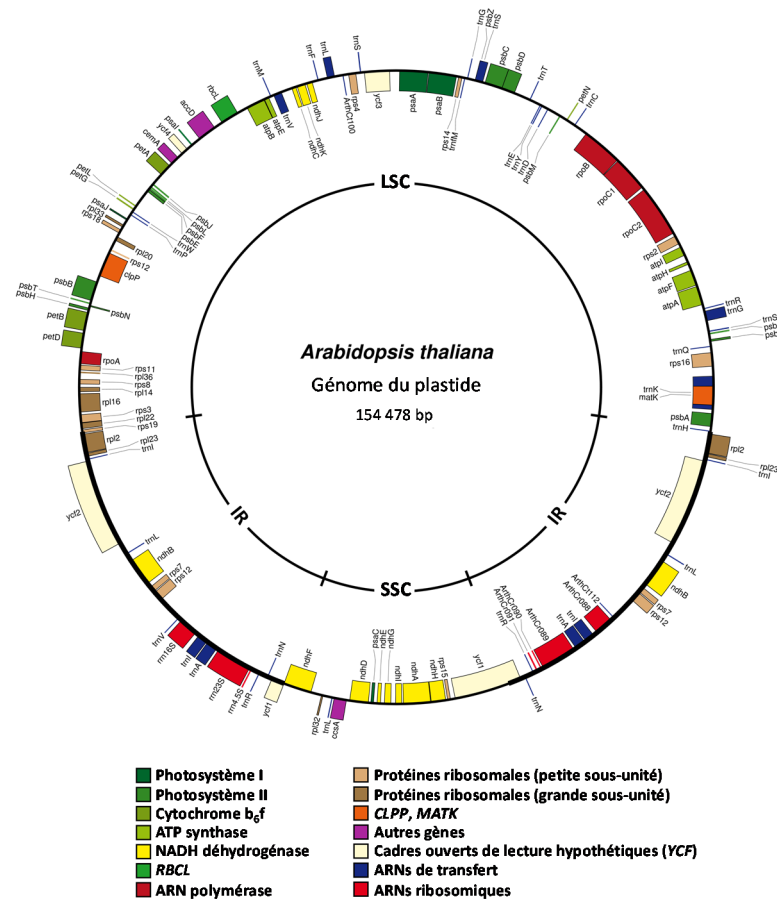


Figure 4. Représentation circulaire du génome du plastide d'*Arabidopsis* et classification fonctionnelle de chacun des gènes. Cette figure a été générée par *Organellar Genome DRAW* [32].

1.3.2 Le nucléoïde du plastide

Les plastides sont des organites hautement polyploïdes, un seul chloroplaste pouvant contenir près de cinquante copies du génome [revu dans 16]. De plus, puisque les cellules végétales contiennent chacune de nombreux plastides, le nombre de copies de ce génome surpasse considérablement le nombre de compléments nucléaires. En effet, chacune des cellules du mésophylle, le principal tissu photosynthétique, peut contenir jusqu'à cent chloroplastes [16]. Le nombre de copies du génome du plastide est d'ailleurs évalué à près de 1000 molécules par cellule végétale [33]. Cette polyploïdie a d'importantes conséquences sur les formes du génome du plastide retrouvées *in vivo*. Ainsi, des formes génomiques distinctes

peuvent être retrouvées au sein d'une même plante. Cette hétéroplasmie du génome du plastide est toutefois compensée par d'efficaces mécanismes de recombinaison et de conversion génique, ceux-ci favorisant généralement la conservation de la forme génomique non-mutée [22].

Curieusement, deux isomères génomiques, qui se distinguent par l'orientation relative des deux régions à simple copie, composent l'ADNpt retrouvé *in vivo* [34]. Lors de cette découverte, il a initialement été proposé que des événements de recombinaison intramoléculaire entre les deux IRs d'une molécule génomique circulaire soient responsables de l'interconversion entre ces deux isomères [34]. Des observations récentes suggèrent cependant que des molécules linéaires et branchées composent la grande majorité des molécules d'ADNpt et que le processus de réplication dépendant de la recombinaison est responsable de ces isomères et des branchements entre des molécules d'ADN linéaires [35] (voir section 1.6.3.3). Ces complexes molécules branchées d'ADNpt semblent d'ailleurs responsables de l'apparition de bandes représentant des oligomères du génome du plastide par électrophorèse en champ pulsé (*PFGE*, « *pulse-field gel electrophoresis*), une technique permettant la séparation de chromosomes sur gel [35-37]. Malgré que les implications de cette découverte restent peu comprises, il est généralement accepté qu'un mélange de molécules circulaires, linéaires et branchées compose l'ADNpt, et que ce mélange hétérogène forme d'énormes complexes [3,23,35].

Ces énormes complexes d'ADNpt sont associés à de nombreuses protéines et molécules d'ARN, et forment, par analogie aux nucléoïdes bactériens, les nucléoïdes du plastide. Bien qu'ils soient retrouvés fixés du côté stromal de la membrane interne dans les proplastides, ces nucléoïdes sont généralement relocalisés vers la face stromale des thylakoïdes lors de la différenciation en chloroplastes [38]. Les nucléoïdes sont ultérieurement séparés en quantité presque égale entre les chloroplastes fils lors du processus de scission des chloroplastes matures [39,40]. Plusieurs études ont analysé la composition du protéome des nucléoïdes par spectrométrie de masse [revu dans 41]. Celles-ci ont révélé qu'en plus de protéines impliquées dans le métabolisme de l'ADNpt, certaines protéines présentes dans les

nucléoïdes participent à l'expression génique et à la maturation des ARNm [42]. Des protéines ribosomales et des protéines accessoires de la traduction ont aussi été détectées dans les nucléoïdes, suggérant que la traduction se produirait de manière co-transcriptionnelle comme chez les procaryotes [42]. Parmi les autres constituants du nucléoïde révélés par ces études sont aussi retrouvées des ADN polymérases assurant la réplication [43], une recombinaise de type RecA impliquée dans la recombinaison [44] et des gyrases permettant la relaxation de l'ADN [45]. Des protéines Whirly stabilisant l'ADN simple-brin (ADNss) [27], et compactant l'ADN [46], sont également associées à l'ADNpt [42]. Enfin, des superoxydes dismutases et des thiorédoxines sont retrouvées en association avec les nucléoïdes et détoxifient les espèces oxygénées réactives (ROS) [42,47,48] (voir Section 1.6.1.2).

1.4 Expression génique dans le plastide

1.4.1 La machinerie de transcription

Bien que le génome du plastide n'encode qu'un petit nombre de gènes, l'expression génique de ce compartiment n'en est pas moins finement régulée, tant au niveau transcriptionnel que traductionnel. Tel que mentionné à la section 1.2.2, deux types d'ARN polymérase, la PEP et la NEP, sont responsables de la transcription des gènes du plastide. La PEP ou « Plastid-encoded RNA polymerase » est similaire à l'ARN polymérase de *E. coli* et est formée de cinq sous-unités. Les quatre premières forment le noyau central de la polymérase et sont encodées au plastide par les gènes *rpoA*, *rpoB*, *rpoC1* et *rpoC2*, ces trois derniers gènes étant transcrits en un seul opéron [49]. La cinquième sous-unité est, quant à elle, encodée au noyau et correspond à chacun des six facteurs sigma (SIG1-6) similaires aux facteurs sigma⁷⁰ bactériens [revu dans 18]. Par leur association aux quatre autres sous-unités, chacun de ces facteurs sigma confère à la PEP une spécificité aux promoteurs de certains gènes du plastide. En conséquence, l'expression génique du plastide se retrouve partiellement sous contrôle nucléaire.

La NEP ou « Nuclear-encoded RNA polymerase » constitue le second type d'ARN polymérase retrouvée dans le plastide et est, comme son nom l'indique, encodée au noyau et ciblée au plastide suite à son expression. Comparativement à la PEP, la NEP est constituée d'une seule sous-unité et est similaires à l'ARN polymérase du bactériophage T7 [50]. Il existe toutefois deux gènes nucléaires pour la NEP : *RPOTp*, qui encode une NEP ciblée qu'au plastide, et *RPOTpm*, encodant une NEP ciblée à la fois au plastide et à la mitochondrie [50,51].

Les gènes du plastide ont été initialement classés selon le type d'ARN polymérase responsable de leur transcription (la NEP, la PEP ou les deux polymérases à la fois). Toutefois, il a été récemment démontré que la quasi-totalité des gènes du plastide contient à la fois un promoteur spécifique à la PEP et un spécifique à la NEP [52]. Quoi qu'il en soit, il est généralement suggéré que la NEP est nécessaire tôt dans le développement des chloroplastes pour l'expression des gènes encodant la PEP [revu dans 16]. La PEP nouvellement formée permet ensuite l'expression des gènes nécessaires à la photosynthèse. Suite à leur transcription, les ARNm du plastide subissent de nombreuses modifications telles que l'épissage, l'édition et le clivage d'ARNs polycistroniques en ARNs monocistroniques [16,53].

1.4.2 La machinerie de traduction

Même si plus d'une ARN polymérase et plusieurs facteurs sigma régulent l'expression génique du plastide au niveau transcriptionnel, la régulation de la traduction constitue pourtant le principal niveau de régulation génique de cet organite [revu dans 54,55]. En effet, de nombreux facteurs d'initiation et d'élongation de la traduction régulent l'expression génique dans le plastide [56-59]. Par ailleurs, des protéines contenant des répétitions pentatricopeptides (PPR) participent également à la régulation de la traduction dans les organites par leur liaison aux molécules d'ARN [revu dans 55]. Avec plus de quatre cents membres chez la plupart des espèces de plantes, les protéines PPR sont impliquées dans une multitude de processus touchant les ARNs, dont, notamment, les étapes de maturation citées ci-haut, la régulation de la traduction de même que la stabilisation et la dégradation des ARN messagers [55].

Un ribosome très similaire aux ribosomes bactériens effectue la traduction dans le plastide [revu dans 54]. Tous les ARNs composant ce ribosome sont encodés dans le plastide et transcrits en un opéron unique qui, suite à des étapes de maturation, génère quatre ARNs ribosomiques matures homologues aux ARNs ribosomiques bactériens [revu dans 54]. Quant aux protéines composant le ribosome du plastide, celles-ci sont encodées par des gènes distribués à la fois dans le génome du plastide et du noyau. La composition protéique du ribosome du plastide est aussi très similaire à celle des ribosomes bactériens, 52 des 57 protéines ribosomales ayant des orthologues chez *E. coli* [revu dans 54]. Afin d'effectuer la traduction, le ribosome du plastide utilise un assortiment d'ARNs de transfert entièrement encodé dans le génome du plastide. Étonnamment, de nombreuses plantes n'encodent que 30 ARNs de transfert différents dans le génome de leurs plastides [revu dans 54]. Pour le moment, aucune évidence ne suggère l'import au chloroplaste d'ARN de transfert encodés au noyau. Il a néanmoins été démontré que ce nombre restreint d'ARNs de transfert est suffisant, un appariement de type « wobble » de la troisième base de plusieurs codons assurant la traduction de la totalité du code génétique [60].

1.5 La réplication de l'ADN du plastide

1.5.1 La machinerie de réplication

Tel que vu à la section 1.3.2, le génome du plastide est présent en de très nombreuses copies dans les cellules végétales et représente près de 25% de la masse totale de l'ADN dans une cellule d'*Arabidopsis* [61]. Une machinerie de réplication de l'ADN est donc dédiée au plastide et permet de maintenir de hauts niveaux d'ADN_{pt} durant la croissance des tissus [33]. L'existence d'une machinerie de réplication de l'ADN_{pt} a d'ailleurs été démontrée, il y a maintenant plus de 30 ans, par purification d'une activité de polymérisation de l'ADN à partir d'extraits chloroplastiques d'épinards [62]. La caractérisation des propriétés biochimiques de cette activité de polymérisation d'ADN suggérait qu'une ADN polymérase distincte des ADN polymérases nucléaires était responsable de la réplication de l'ADN du plastide [62]. Il a depuis été démontré que deux ADN polymérases hautement similaires à l'ADN polymérase I

des cyanobactéries sont encodées dans le génome nucléaire et ciblées, suite à leur expression, dans les plastides du riz [63,64], d'*Arabidopsis* [64,65] et du tabac [66]. Il a aussi été remarqué que l'expression d'une seule de ces deux ADN polymérases, *POLIB*, est augmentée chez *Arabidopsis* suite à une exposition au peroxyde d'hydrogène, suggérant un rôle plus spécifique dans la réparation de l'ADN pour cette polymérase [64].

Chez *Arabidopsis*, une caractérisation *in vivo* de mutants des gènes *POLIA* et de *POLIB* a établi la participation de chacune de ces polymérases dans la réplication de l'ADN des organites [43]. En effet, les mutations *polla* et *pollb* provoquent une diminution des niveaux d'ADN mitochondrial et chloroplastique. Le rôle essentiel, mais redondant, de *POLIA* et *POLIB* dans la réplication de l'ADN des organites a de plus été démontré par la létalité associée à la combinaison des mutations homozygotes de ces deux gènes [43]. Outre la redondance de ces deux polymérases dans le processus de réplication, cette étude a aussi confirmé la spécialisation de *POLIB* dans la réparation de l'ADN de la mitochondrie et du plastide [43] (voir section 1.6.3.4).

Comparativement à ces deux ADN polymérases, les autres acteurs participant à la réplication de l'ADN du plastide demeurent très peu caractérisés. Néanmoins, des études ont démontré qu'il est possible d'isoler, par purification biochimique, des activités d'ADN primase, d'ADN hélicase et d'ADN gyrase à partir d'extraits de chloroplastes de pois [67-69]. Deux homologues de l'hélicase bactérienne RecQ ont d'ailleurs été identifiés dans les chloroplastes du riz [70]. D'autre part, le génome d'*Arabidopsis* encode des homologues de chacune des sous-unités A et B des ADN gyrases procaryotes [45]. Celles-ci sont ciblées à la fois au plastide et à la mitochondrie et, de manière similaire à leurs homologues bactériens, catalysent la relaxation des surenroulements positifs de l'ADN causés par la réplication [45,71].

1.5.2 Les mécanismes de réplication

Puisque le génome du plastide est retrouvé sous forme d'un mélange de molécules circulaires, linéaires et branchées, plusieurs mécanismes participent donc à sa réplication [revu dans 72]. Trois principaux mécanismes ont été proposés jusqu'ici : le modèle de la

boucle D double (*Double D-Loop*), le modèle du cercle roulant (*Rolling circle*) et le modèle de réplication dépendant de la recombinaison. Tout comme la proportion de chacune des formes génomiques citées plus tôt, la contribution de chacun de ces mécanismes à la réplication de l'ADNpt reste un sujet de controverse. Toutefois, alors que les modèles de la boucle D double et du cercle roulant ne sont compatibles qu'avec des molécules génomiques de forme circulaire, le modèle de réplication dépendant de la recombinaison nécessite quant à lui une molécule d'ADN linéaire en plus d'une seconde molécule, circulaire ou linéaire [revu dans 72].

Le modèle de la boucle D double, ou boucle de déplacement, s'amorcerait par l'initiation de la synthèse d'ADN à l'emplacement de certaines séquences spécifiques du génome, nommées origines de réplication. Même si plusieurs contradictions existent quant à l'emplacement des origines de réplication du plastide, il semble néanmoins probable que deux origines d'orientation opposée soient localisées au sein de chacun des IRs [revu dans 72]. L'initiation de la synthèse d'ADN sur ces deux origines de réplication produirait deux boucles D d'orientation convergente, dans lesquelles un court segment d'ADN nouvellement synthétisé maintiendrait séparés les deux brins d'ADN matrices (Fig. 5A). De telles structures ont d'ailleurs été observées par microscopie électronique dans l'ADN chloroplastique du pois et du maïs [73]. Suite à l'élongation de ces deux brins d'ADN nouvellement synthétisés, la rencontre des deux boucles D générerait une bulle de réplication bidirectionnelle. Alors que chacune des boucles D dépasserait les origines de réplication, une synthèse discontinue d'ADN s'amorcerait sur les brins opposés. La rencontre des deux fourches de réplication, à près de 180 degrés des origines de réplication, complèterait finalement la réplication de la molécule génomique circulaire.

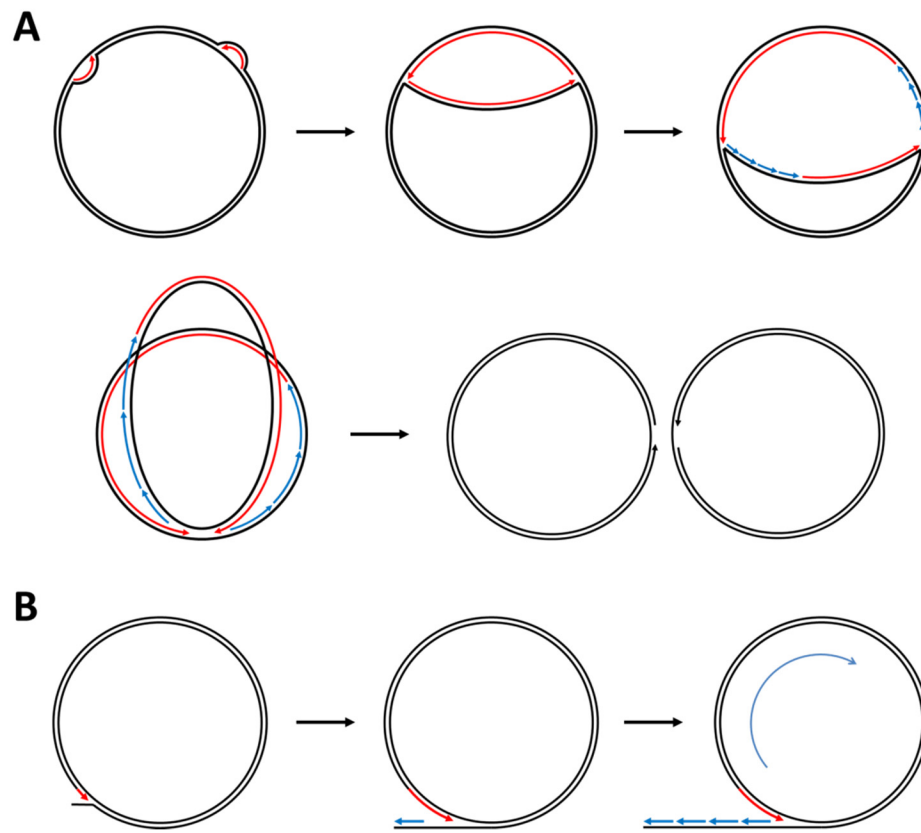


Figure 5. Représentation schématique des modèles de la boucle D double et du cercle roulant. (A) Modèle de la boucle D double. (B) Modèle du cercle roulant. Les traits noirs représentent les brins matrices. Les traits rouges et bleus représentent, respectivement, les brins nouvellement synthétisés continus et discontinus. Les flèches représentent la direction de la synthèse d'ADN. Cette figure est adaptée de [74].

Quant au modèle du cercle roulant, il est habituellement suggéré que celui-ci s'initierait au site de terminaison de la réplication par le mécanisme de la boucle D double. [72,75]. Selon ce modèle, suite à l'achèvement d'une ronde de réplication par boucle D double, l'extrémité 3' nouvellement synthétisée déplacerait sa propre extrémité 5' afin de poursuivre son élongation. Une synthèse discontinue d'ADN s'initierait alors sur le brin déplacé (Fig. 5B). Toujours selon ce modèle, puisque l'élongation du brin continu peut effectuer plusieurs tours de la molécule génomique circulaire, ce type de réplication pourrait mener à la formation de concatémères tête-à-queue du génome.

Tel que mentionné précédemment, de nombreuses contradictions existent quant aux positions exactes des origines de réplication du plastide et quant à leur contribution réelle dans le processus de réplication. Chez le tabac, l'inactivation de trois des quatre origines de réplication du génome du plastide n'affecte que très peu sa capacité à se répliquer [76]. De plus, chez l'algue verte *C. reinhardtii*, l'inhibition de la réplication du plastide par la novobiocine, un inhibiteur des gyrases, déclenche un type de réplication dépendant de la recombinaison [77]. Un mécanisme alternatif de réplication de l'ADN_{pt} dépendant de la recombinaison a d'ailleurs été récemment proposé [35,72]. Selon ce modèle, l'extrémité 3' d'une molécule linéaire d'ADN envahirait un duplex d'ADN homologue et mènerait à l'établissement d'une fourche de réplication (voir section 1.6.3.3). Cet envahissement d'un duplex d'ADN matrice par une molécule linéaire expliquerait les nombreuses molécules branchées composant l'ADN_{pt} des nucléoïdes [35].

1.6 Les mécanismes assurant la stabilité du génome du plastide

Même s'ils sont peu nombreux, les gènes encodés dans le plastide servent presque exclusivement au processus de photosynthèse et sont, conséquemment, absolument essentiels au métabolisme énergétique des végétaux. Le maintien de la stabilité de l'ADN_{pt} est donc primordial et doit être assuré par une multitude de processus prévenant les dommages à l'ADN et permettant, le cas échéant, une réparation adéquate de l'ADN endommagé. Ces mécanismes semblent d'ailleurs particulièrement efficaces, le génome du plastide présentant un taux de mutations plus faible que l'ADN nucléaire [20]. Dans cette prochaine section sont discutées les principales menaces à l'intégrité génomique du plastide dont, notamment, les déséquilibres photosynthétiques et les ROS qui leur sont associés, les rayons ultraviolets (UV) et les brisures double-brins de l'hélice d'ADN (*Double-Strand Breaks*, DSB). Les moyens employés pour leur faire face sont également discutés.

1.6.1 Les mécanismes de prévention et de réparation des dommages oxydatifs

Bien que les réactions d'oxydoréduction caractérisant la photosynthèse soient à la base du métabolisme énergétique chez les plantes, celles-ci génèrent aussi des ROS qui, à de trop hautes concentrations, mènent à de graves déséquilibres rédox. Par leur fort potentiel oxydant, les ROS peuvent endommager l'ensemble des constituants cellulaires, dont les acides nucléiques, les lipides et les protéines [78]. De nombreuses conditions environnementales, telles que les hautes intensités lumineuses, les extrêmes de température et la sécheresse, stimulent la production de ROS par la PET [79]. Dans de telles conditions de stress abiotique, l'absorption lumineuse excède la capacité de la PET à drainer efficacement les électrons, menant du même coup à une hyper-réduction des composantes de la PET. Des électrons sont alors transférés à des molécules d'oxygène, notamment par le PSI, et génèrent des anions superoxydes (O_2^-) et du peroxyde d'hydrogène (H_2O_2) (Fig. 6-7). D'autre part, l'hyper-réduction du PSII et de ses complexes d'antennes mène à la production d'un type de ROS distinct, l'oxygène singulet (1O_2), par transfert d'énergie entre une molécule de chlorophylle excitée et une molécule d'oxygène [revu dans 80]. En plus des fluctuations environnementales, de hautes concentrations de ROS sont générées lors d'un déséquilibre de la PET. De tels déséquilibres sont induits, notamment, par une stœchiométrie inadéquate entre les différentes composantes de la PET ou par la présence de sous-unités non-fonctionnelles des photosystèmes [81]. Tel que discuté plus en détails dans le texte qui suit, les ROS n'ont pas que des effets délétères. En effet, ceux-ci initient des voies de signalisation permettant aux plantes de faire face au stress oxydatif.

1.6.1.1 Les mécanismes prévenant la production de ROS par la photosynthèse

Le processus de photosynthèse confère aux chloroplastes une perception très sensible des fluctuations environnementales [82]. Afin d'éviter de graves déséquilibres rédox occasionnés par ces fluctuations, de nombreux mécanismes et points de contrôle adaptent l'absorption lumineuse et la capacité photosynthétique de la PET. Ces mécanismes préviennent donc les dommages oxydatifs des constituants cellulaires, tels les acides nucléiques composant le génome du plastide.

A) La réponse à court terme aux déséquilibres photosynthétiques

Parmi les mécanismes conférant une adaptation à court terme, il convient de mentionner le « *Non-photochemical quenching* », par lequel des pigments de la famille des caroténoïdes dissipent sous forme de chaleur l'excès d'énergie lumineuse captée [83]. Les transitions d'état de la PET (*State transitions*) constituent un autre mécanisme permettant d'éviter un déséquilibre rédox photosynthétique lors de conditions d'absorption lumineuse inégales entre les deux photosystèmes [revu dans 84]. Dans ce processus, l'hyper-réduction du pool de plastoquinone, l'accepteur d'électrons du PSII, stimule chez *Arabidopsis* l'activité de la protéine kinase des thylakoïdes STN7 [85,86]. Une fois activée, la kinase STN7 phosphoryle le complexe d'antenne mobile LHCII (*Light-harvesting complex II*), provoquant la relocalisation de celui-ci du PSII au PSI [84]. Cette relocalisation de LHCII redistribue la capacité d'absorption lumineuse entre les deux photosystèmes et prévient, à court terme, un déséquilibre de la PET et la production de ROS qui y est associée.

B) La réponse à long terme aux déséquilibres photosynthétiques

Afin de s'adapter aux fluctuations environnementales, les chloroplastes modifient aussi de façon plus durable la composition protéique de la PET et adaptent l'expression de nombreux gènes participant à la réponse au stress oxydatif. Ainsi, par des modifications des transcriptomes chloroplastiques et nucléaires, les chloroplastes ajustent la stœchiométrie des deux photosystèmes [revu dans 87], optimisent les niveaux d'expression des complexes d'antennes [85,88] et accumulent de hauts niveaux de protéines anti-oxydantes [89]. Des voies de signalisation rétrograde initiées au chloroplaste sont à l'origine des ajustements de l'expression génique nucléaire. Deux principales classes de signaux initient ce type de signalisation en conditions de stress photosynthétique : l'état rédox des composantes de la PET et l'accumulation de certains métabolites [revu dans 87,90].

La signalisation rétrograde dépendante de l'état rédox de la PET

La protéine kinase STN7 est à la base des voies de signalisation rétrograde dépendantes de l'état rédox de la PET [91]. Lors de l'hyper-réduction du pool de plastoquinone, STN7 initie

une cascade de signalisation rétrograde en phosphorylant un substrat encore inconnu [87,92]. Il en résulte une modification des transcriptomes nucléaire et chloroplastique, et un ajustement de la composition de la PET. Une forte réduction du pool de plastoquinone mène d'ailleurs à une modification de l'expression de 937 gènes nucléaires chez *Arabidopsis* [93]. De manière intéressante, 85 % de ces gènes sont insensibles à l'état rédox de la plastoquinone chez un mutant *stn7* [93]. Cette kinase semble donc jouer un rôle primordial dans l'adaptation de la capacité photosynthétique. Par ailleurs, certaines évidences suggèrent que l'état rédox du PSI contribuerait également à ce type de signalisation rétrograde [94].

La signalisation rétrograde dépendante de l'accumulation de métabolites

Le déséquilibre rédox de la PET mène aussi à l'accumulation de certains métabolites déclenchant des voies de signalisation rétrograde dans le chloroplaste. Parmi ces métabolites, les ROS sont sans contredit les mieux caractérisés, l'étude du mutant *flu* d'*Arabidopsis* y ayant largement contribué [95,96]. Ce mutant cause l'accumulation d'un précurseur de la chlorophylle, nommé protochlorophyllide, dans les chloroplastes à la noirceur, ce qui génère une grande quantité d'oxygène singulet lors de la ré-exposition des tissus à la lumière [96]. Dans le mutant *flu*, cette haute concentration d'oxygène singulet est associée à un arrêt de croissance, à l'apparition de lésions et au blanchiment des tissus. Curieusement, ce phénotype ne semble pourtant pas être une conséquence de la toxicité de l'oxygène singulet. En effet, les mutations d'EXECUTER1 et 2, deux protéines impliquées dans la signalisation rétrograde, préviennent l'apparition de ce phénotype [97,98]. Par conséquent, il paraît davantage probable que le phénotype associé au mutant *flu* résulte d'une réponse génétique nucléaire programmée induite par signalisation rétrograde de l'oxygène singulet. La pertinence d'une telle réponse génétique programmée dans le contexte d'une plante de type sauvage est toutefois contestable, celle-ci ne favorisant pas la survie de l'individu en condition de stress. Cependant, il a été récemment démontré que cette signalisation rétrograde est nécessaire à l'acclimatation à un stress environnemental modéré chez des plantes de type sauvage [99].

Outre l'oxygène singulet, il a aussi été démontré que le peroxyde d'hydrogène chloroplastique est à l'origine d'un signal rétrograde régulant un grand nombre de gènes nucléaires [100]. Les voies de signalisation dépendantes du peroxyde d'hydrogène et de l'oxygène singulet ne sont pourtant pas indépendantes l'une de l'autre, celles-ci ayant des effets antagonistes sur la réponse nucléaire [101]. Dans le contexte d'une plante de type sauvage, les effets de l'oxygène singulet paraissent donc être contrebalancés par ceux du peroxyde d'hydrogène afin d'ajuster très finement la réponse génétique nucléaire et faire face aux stress oxydatifs [80,101].

Parmi les autres métabolites s'accumulant en conditions de déséquilibre photosynthétique, il convient de mentionner le MEcPP (méthylérythritol cyclodiphosphate) [102], un intermédiaire de la voie de biosynthèse chloroplastique des isoprénoïdes, le PAP (3'-phosphoadenosine 5'-phosphate) [103], un nucléotide modifié servant à la biosynthèse de composés sulfatés et le β -cyclocitral [104], un produit d'oxydation du caroténoïde β -carotène par l'oxygène singulet (Fig. 6). Comme pour les ROS, une reprogrammation de l'expression génique nucléaire est associée à de hautes concentrations chloroplastiques de chacun de ces métabolites [revu dans 90]. La reprogrammation génétique nucléaire associée au β -cyclocitral est d'ailleurs très similaire à celle du mutant *flu* [104]. Cette similitude suggère que les produits d'oxydation des caroténoïdes, la première ligne de défense chloroplastique contre l'oxygène singulet, constituent un des signaux initiant les voies de signalisation rétrograde dépendantes de ce type de ROS [104]. La haute volatilité du β -cyclocitral et sa capacité de diffusion au cytosol supportent d'autant plus cette hypothèse. D'autre part, l'exposition au β -cyclocitral favorise une résistance au stress oxydatif photosynthétique chez *Arabidopsis*, et renforce l'idée que ce composé sert de molécule de signalisation [104]. L'interrelation possible entre le β -cyclocitral et les protéines EXECUTERS pour la signalisation de l'oxygène singulet reste, pour l'instant, inconnue.

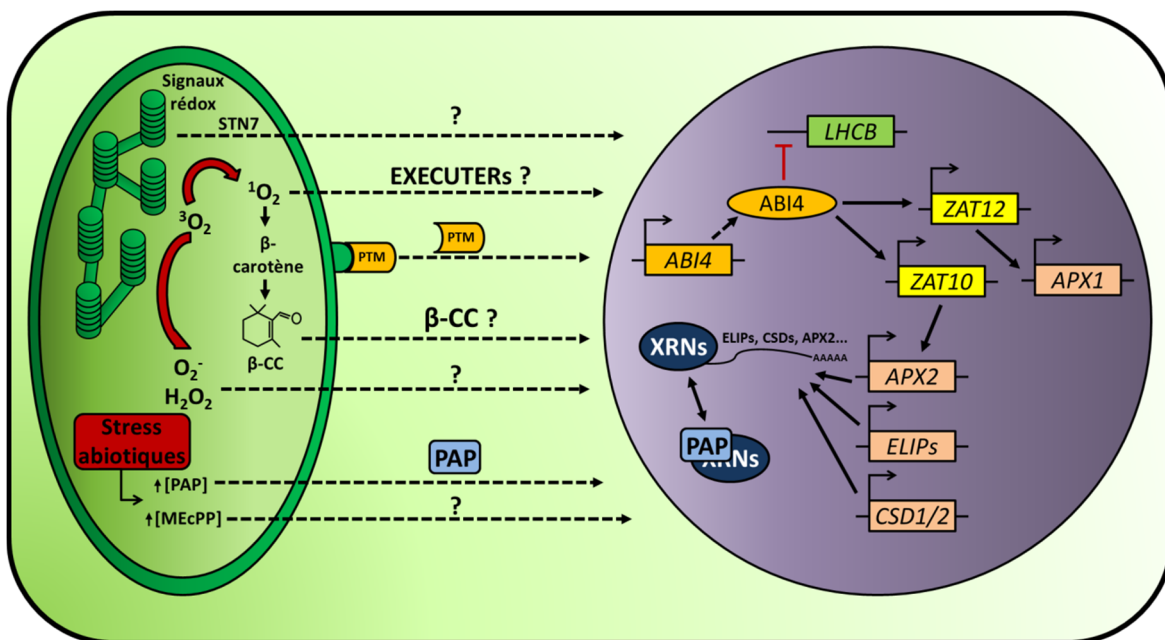


Figure 6. Les voies de signalisation rétrograde chloroplaste-noyau et quelques-uns de leurs principaux effets sur la reprogrammation génétique nucléaire.

Le β -CC et le PAP ont probablement la capacité de diffuser au noyau afin de moduler l'expression génique nucléaire. Lorsqu'en concentration suffisante dans le noyau, le PAP aurait la capacité d'inhiber des XRNs, menant à l'accumulation de plusieurs transcrits, tels que ceux encodant les ELIPs, APX2, CSD1 et CSD2. Suite à son clivage protéolytique, le facteur de transcription PTM diffuse vers le noyau et active l'expression du facteur de transcription ABI4. ABI4 réprime alors l'expression de gènes associés à la photosynthèse tels que *LHCB*, mais active l'expression des facteurs de transcription ZAT10 et 12. Le mécanisme employé pour acheminer au noyau la plupart des autres voies de signalisation rétrograde reste inconnu pour l'instant. $^3\text{O}_2$: dioxygène; $^1\text{O}_2$: oxygène singulet; O_2^- : superoxyde; H_2O_2 : peroxyde d'hydrogène; β -CC : β -cyclocitral; XRNs : exoribonucléases.

De manière similaire au β -cyclocitral, des évidences suggèrent que le PAP possède la capacité de diffuser du chloroplaste vers le noyau [103]. On ne sait toutefois pas si le MEcPP a cette capacité. Pour le moment, les mécanismes permettant de relayer ces signaux rétrogrades au noyau ne sont donc que partiellement compris. Néanmoins, une récente découverte a révélé l'implication du facteur de transcription PTM (*PHD type transcription factor with transmembrane domains*) dans l'acheminement des signaux rétrogrades au noyau. Le facteur de transcription PTM est lié au côté cytoplasmique de la membrane chloroplastique externe

mais, en réponse à de nombreuses voies de signalisation rétrograde, son domaine N-terminal est clivé et diffuse au noyau afin de moduler l'expression de nombreux gènes participant à la réponse au stress [105] (Fig. 6).

C) La réponse nucléaire aux voies de signalisation rétrograde

Une fois acheminée au noyau, ces différentes voies rétrogrades déclenchent des réponses génétiques distinctes, mais souvent très similaires, par la modulation de l'expression de nombreux gènes [106]. Plusieurs voies de signalisation rétrograde convergent d'ailleurs vers quelques régulateurs transcriptionnels principaux, dont les facteurs de transcription GLK1, GLK2 et ABI4 [revu dans 107]. Alors que le facteur de transcription ABI4 ralentit la différenciation des proplastides en chloroplastes, les facteurs GLK1 et GLK2 en stimulent, quant à eux, la biogénèse [108]. Pour cette raison, les niveaux cellulaires de GLK1 et GLK2 diminuent lors d'un stress photosynthétique [109]. Au contraire des facteurs GLK, l'expression d'ABI4 augmente dans de telles conditions, le facteur de transcription PTM activant directement sa transcription [105] (voir section précédente). Par l'augmentation des niveaux d'ABI4, de nombreux gènes associés à la photosynthèse sont réprimés, dont certaines composantes des complexes d'antennes [110]. D'autre part, de manière à développer une tolérance au stress photosynthétique, ABI4 active l'expression des facteurs de transcription ZAT10 et ZAT12 [110], d'importants régulateurs transcriptionnels opérant dans l'adaptation au stress oxydatif et aux hautes intensités lumineuses [111,112]. Les facteurs de transcription ZAT12 et ZAT10 stimulent notamment l'expression des ascorbate peroxydases 1 et 2 (APX1 et APX2), respectivement, d'importantes enzymes de détoxification du peroxyde d'hydrogène [112,113].

Toutefois, les voies de signalisation rétrograde ne contrôlent pas l'expression génique qu'au niveau de la régulation transcriptionnelle. En effet, lors des fluctuations lumineuses, certaines voies de signalisation rétrograde modulent l'expression génique nucléaire au niveau post-transcriptionnel par des mécanismes d'épissage alternatif et de dégradation des ARNm [103,114]. L'augmentation de la concentration cellulaire de PAP semble d'ailleurs

inhiber l'activité d'exoribonucléases régulant l'expression de nombreux gènes de réponse aux stress [103] (Fig. 6). Chez *Arabidopsis*, ces exoribonucléases régulent l'expression des facteurs de transcription ZAT10 et ZAT12, de certaines enzymes de détoxification des ROS dont l'ascorbate peroxydase 2 (APX2) et les superoxydes dismutases CSD1 et CSD2, et des protéines ELIP1 et ELIP2 (*early light-induced proteins*) [103,115]. Comme leur nom l'indique, les protéines ELIPs sont rapidement surexprimées lors d'un transfert à haute intensité lumineuse ou lors d'un stress oxydatif [116,117]. De plus, ces protéines sont structurellement similaires aux complexes d'antennes LHC [118,119] et, tout comme ceux-ci, ont la capacité de lier à la fois des molécules de chlorophylle et de caroténoïde [119,120]. Les protéines ELIPs ne servent cependant pas à la capture de photons. Il est plutôt suggéré que ces protéines empêchent l'accumulation de molécules de chlorophylle libres dans le chloroplaste, celles-ci pouvant générer de l'oxygène singulet et, conséquemment, un important stress oxydatif [116,121]. Par leur liaison aux ELIPs, les molécules de chlorophylle libres excitées pourraient en effet transférer leur surplus énergétique aux molécules de caroténoïde situées à proximité, prévenant ainsi la génération d'oxygène singulet [116].

1.6.1.2 La détoxification des ROS au chloroplaste

Même si les voies de signalisation rétrograde limitent la production de ROS par le processus de photosynthèse, les modifications du transcriptome nucléaire nécessitent généralement un délai de plusieurs heures avant d'être complétées. De plus, cette adaptation de l'expression génique nucléaire ne suffit parfois pas à enrayer le stress oxydatif chloroplastique. Il est donc essentiel que des mécanismes assurent la détoxification des ROS dans le chloroplaste. Alors que dans ce compartiment les caroténoïdes constituent la principale défense contre l'oxygène singulet en dissipant l'excès énergétique sous forme de chaleur [revu dans 122], une complexe machinerie enzymatique détoxifie le superoxyde et le peroxyde d'hydrogène [revu dans 78,123]. Une machinerie spécifique est dédiée à la détoxification de ces types de ROS afin d'éviter la formation de radicaux hydroxyles ($\bullet\text{OH}$) hautement réactifs et menaçant l'ensemble des composants cellulaires, dont l'ADN [124]. Le cycle d'Haber-Weiss et la réaction de Fenton peuvent en effet transformer le superoxyde et le peroxyde d'hydrogène

en radicaux hydroxyles par le biais d'une catalyse chimique par le fer [125]. Aucune machinerie enzymatique ne permet la détoxification de ce type de ROS.

Deux principales étapes assurent la détoxification du superoxyde et sa conversion en molécules d'eau : 1) la transformation de deux molécules de superoxyde en peroxyde d'hydrogène par les enzymes de la famille des superoxydes dismutases et 2) la réduction du peroxyde d'hydrogène en eau par une enzyme de type peroxydase. Sept superoxyde dismutases (SODs) sont encodées dans le génome d'*Arabidopsis*, dont trois sont dépendantes du fer (*FSD1*, *FSD2* et *FSD3*), trois autres dépendent du cuivre/zinc (*CSD1*, *CSD2* et *CSD3*) et une est dépendante du manganèse (*MSD1*) [revu dans 126]. Parmi celles-ci, *FSD2*, *FSD3* et *CSD2* sont localisées au chloroplaste [127-129], et participent à la tolérance du stress oxydatif lié à la photosynthèse. En effet, chez *Arabidopsis*, chacune des mutations *fsd2*, *fsd3* et *csd2* cause un développement anormal des chloroplastes en plus d'être associée à un retard de croissance et à une coloration pâle des feuilles [127,130]. Une croissance à faible intensité lumineuse rétablie cependant partiellement la croissance des tissus, démontrant l'importance de ces enzymes pour l'équilibre rédox du chloroplaste. Par ailleurs, *CSD2* est principalement retrouvée du côté stromal des thylakoïdes et, conséquemment, à proximité de l'ADNpt et des sites de production de ROS par la PET [129]. De manière similaire à *CSD2*, *FSD3* est localisée spécifiquement aux nucléoïdes [127]. Quant à *FSD2*, celle-ci est retrouvée plus dispersée dans le stroma du chloroplaste [127]. *FSD2* et *FSD3* interagissent toutefois physiquement, suggérant que *FSD3* est capable de recruter *FSD2* aux nucléoïdes [127]. Étonnamment, une étude récente a démontré que *FSD2* et *FSD3* sont des composantes essentielles du complexe transcriptionnel de la PEP [131], appuyant la thèse de l'importance des superoxydes dismutases pour la stabilité de l'ADNpt et son expression.

Afin d'éviter la diffusion du peroxyde d'hydrogène dans le stroma du chloroplaste et sa transformation en radical hydroxyle, les superoxydes dismutases sont retrouvées à proximité de la machinerie convertissant le peroxyde d'hydrogène en eau [revu dans 126]. Trois types de peroxydases permettent cette conversion dans le chloroplaste : les peroxydases à ascorbate (APX), les peroxydases à glutathion (GPX) ainsi que les peroxyrédoxines (PRX). Les

APXs et GPXs utilisent l'ascorbate (AsA) et le glutathion (GSH), respectivement, comme donneur d'électrons lors de la réduction du peroxyde d'hydrogène en eau [126,132] (Fig. 7). Suite à cette catalyse, l'ascorbate se retrouve sous forme oxydée, nommée monodihydroascorbate (MDA) et le glutathion sous forme de disulfure de glutathion (GSSG). De manière à régénérer les pools d'AsA et de GSH, les enzymes MDA réductase (MDAR) et glutathion réductase (GR) utilisent le NAD(P)H pour réduire le MDA et le GSSG. Des isoformes de chacune de ces enzymes sont d'ailleurs retrouvées dans le chloroplaste [133-140]. De plus, deux isoformes de l'APX chloroplastique sont retrouvées chez la plupart des plantes, une APX étant liée à la membrane des thylakoïdes (tAPX) et l'autre étant stromale et soluble (sAPX) [133,141]. Bien que ces deux APXs soient nécessaires à la détoxification efficace du peroxyde d'hydrogène [142], la perte de tAPX semble avoir un impact plus important que celle de sAPX chez *Arabidopsis* [134]. Cette observation démontre probablement l'importance du couplage entre les SODs et les APXs à l'endroit de la membrane des thylakoïdes.

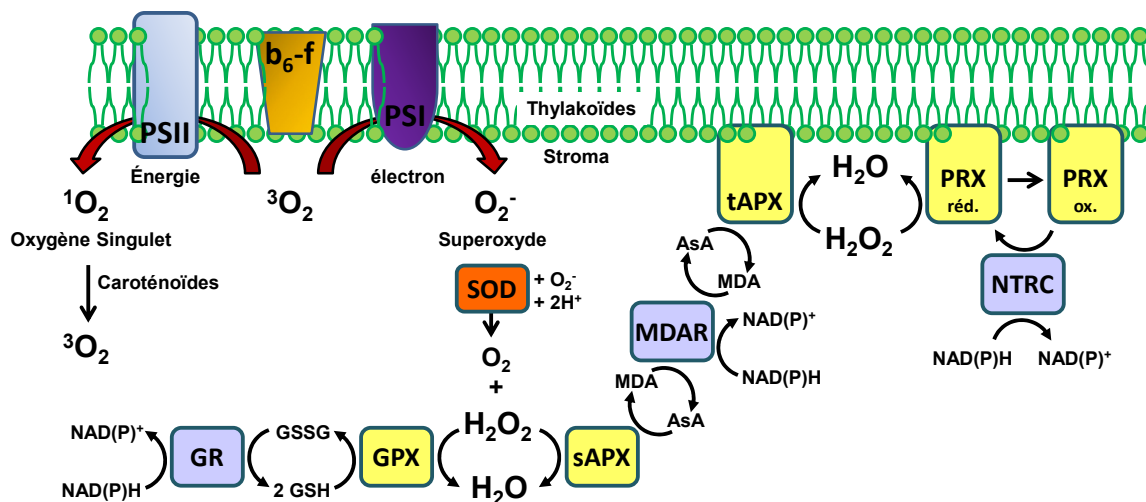


Figure 7. Les mécanismes et enzymes participant à la détoxification des ROS dans le chloroplaste. Les caroténoïdes constituent la principale défense contre l'oxygène singulet en dissipant l'excès énergétique sous forme de chaleur. Des enzymes de la famille des superoxydes dismutases (SODs) transforment le superoxyde en peroxyde d'hydrogène. Le peroxyde d'hydrogène est ensuite transformé en eau par le biais des peroxydases à ascorbate (APX), des peroxydases à glutathion (GPX) et des peroxyrédoxines (PRX). Les pools d'ascorbate (AsA) et de glutathion (GSH) sont régénérés par les enzymes MDA réductase (MDAR) et glutathion réductase (GR) qui utilisent le NAD(P)H comme

pouvoir réducteur. Les PRXs oxydées sont réduites par la thiorédoxine réductase stromale (NTRC) qui utilise aussi le NAD(P)H comme pouvoir réducteur. $^3\text{O}_2$: dioxygène; $^1\text{O}_2$: oxygène singulet; O_2^- : superoxyde; H_2O_2 : peroxyde d'hydrogène; sAPX : peroxydase à ascorbate stromale soluble; tAPX : peroxydase à ascorbate liée à la membrane des thylakoïdes; MDA : monodihydro ascorbate; GSSG : disulfure de glutathion.

Plusieurs familles de PRXs sont retrouvées dans le chloroplaste attachées à la membrane des thylakoïdes [143-145]. Parmi celles-ci, les PRXs à deux cystéines sont les plus étudiées. Les deux cystéines retrouvées au sein de la séquence de ces PRXs réduisent le peroxyde d'hydrogène en eau et, suite à cette catalyse, se retrouvent oxydées par la formation d'un pont disulfure. Le NAD(P)H sert également de pouvoir réducteur pour régénérer la capacité détoxifiante des PRXs. Pour ce faire, la thiorédoxine réductase stromale (NTRC) catalyse la réduction des PRXs par le biais de son domaine thiorédoxine et du NAD(P)H [146-148] (Fig. 7). Il a récemment été démontré que les PRXs chloroplastiques travaillent de concert avec les autres types de peroxydases pour combattre le stress oxydatif. En effet, les mutations de tAPX et des deux PRXs chloroplastiques à deux cystéines ont un effet synergique sur la sensibilité aux hautes intensités lumineuses chez *Arabidopsis* [149].

1.6.1.3 La réparation des dommages oxydatifs de l'ADN chloroplastique

Même s'ils sont très sophistiqués, les mécanismes prévenant les déséquilibres rédox photosynthétiques et l'accumulation de ROS dans le chloroplaste ne sont pas infallibles et des dommages oxydatifs de l'ADNpt surviennent fréquemment. L'oxydation de l'ADN peut mener à la formation de plusieurs dizaines de dérivés de bases azotées [150], dont notamment la 8-oxoguanine et la thymine glycol (Fig. 8). De telles lésions sont hautement mutagènes, celles-ci pouvant bloquer la progression des fourches de réplication [151,152] en plus d'altérer la séquence d'ADN [153]. Les ROS peuvent aussi induire la désamination des résidus cytosines, les convertissant en résidus uraciles [154,155].

Le processus de réparation de l'ADN par excision de bases (*Base excision repair*, BER) permet toutefois de minimiser l'effet mutagène de ces lésions par le remplacement des nucléotides endommagés. Ce mécanisme débute par l'hydrolyse des liens N-glycosidiques des bases

azotées endommagées par une ADN glycosylase, afin de générer un site apurinique/apyrimidique (site AP). Les ADN glycosylases sont généralement classées en deux catégories, selon qu'elles reconnaissent les purines ou les pyrimidines endommagées. Certaines ADN glycosylases, dites bifonctionnelles, ont de plus la capacité de cliver le squelette de l'ADN en 3' du site AP. La suite du processus de réparation dépend néanmoins d'un clivage du squelette en 5' du site AP par une endonucléase, de façon à générer une extrémité 3' libre permettant l'ajout de nucléotides en amont du site endommagé. Le site AP est finalement comblé par l'ajout d'au moins un nucléotide. Les nucléotides résiduels sont ensuite clivés et le squelette de l'ADN est religué [revu dans 156].

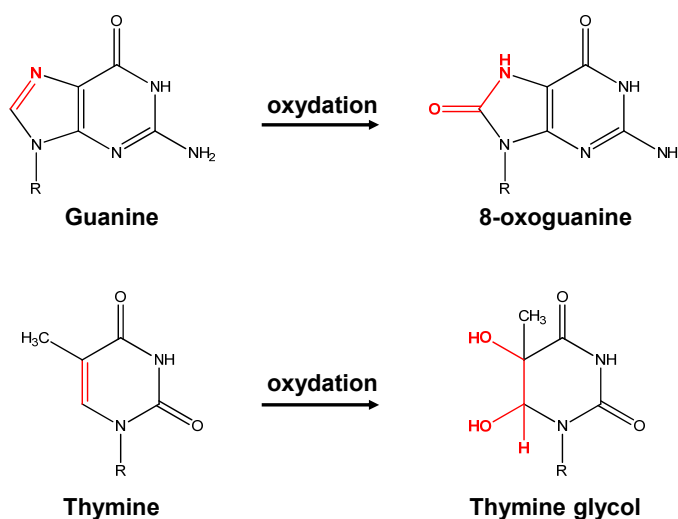


Figure 8. L'oxydation des bases azotées génère de nombreux dérivés de bases azotées, dont la 8-oxoguanine et la thymine glycol. Cette figure est adaptée de [74].

Bien que l'existence de ce mécanisme de réparation dans le chloroplaste soit relativement bien établie, seuls quelques acteurs de cette machinerie de réparation ont été identifiés jusqu'à présent. Ainsi, dans le chloroplaste d'*Arabidopsis*, les ADN glycosylases bifonctionnelles NTH1 et NTH2 excisent les pyrimidines endommagées et clivent le squelette de l'ADN [157,158]. L'existence d'une ADN glycosylase chloroplastique permettant l'excision des purines endommagées est cependant très incertaine. Étonnamment, les plantes sont pourtant les seuls eucaryotes à encoder un homologue de l'ADN glycosylase bactérienne

MUTM, excisant les purines oxydées [159]. Il semble donc probable que cette glycosylase ait été acquise au moment de l'endosymbiose du plastide. De plus, plusieurs isoformes de ce gène sont produites par épissage alternatif et seules certaines d'entre elles possèdent une séquence de localisation nucléaire, suggérant que certaines isoformes pourraient être ciblées au chloroplaste [160]. La présence de MUTM au chloroplaste reste toutefois incertaine. D'autre part, la purification d'une activité ADN glycosylase ciblant la désoxyuridine a été observée dans les chloroplastes du maïs [161]. Finalement, l'endonucléase ARP est probablement chargée du clivage du squelette d'ADN en 5' des sites AP dans le chloroplaste d'*Arabidopsis* [158,162].

1.6.2 Les mécanismes de réparation des dommages à l'ADN induits par les rayons UV

L'exposition à la lumière est absolument nécessaire au processus de photosynthèse. Néanmoins, celle-ci contraint les végétaux à s'exposer continuellement aux rayons UV-B (280-320 nm) pouvant altérer les bases de l'ADN et mener à la formation de liens covalents entre pyrimidines adjacentes. Les rayons UV-B induisent ainsi deux principales lésions dans l'ADN : les dimères cyclobutaniques de pyrimidines (*Cyclobutane pyrimidine dimer*, CPD) et les pyrimidines (6-4) pyrimidones (photo-produits 6-4) (Fig. 9). De manière similaire aux lésions de l'ADN induites par les ROS, les CPDs et les photo-produits 6-4 sont hautement mutagènes et bloquent la progression des ARN et ADN polymérases [163,164]. Afin de réparer ces lésions, les enzymes de la famille des photolyases captent l'énergie de la lumière bleue et catalysent la séparation des dimères de bases azotées en monomères. Chez de nombreux organismes, cette photo-réparation est complétée par le mécanisme d'excision de nucléotides (*Nucleotide-excision repair*). Même si ce mécanisme répare les lésions induites par l'UV-B dans le noyau d'*Arabidopsis* [165], son existence demeure incertaine dans le chloroplaste [revu dans 166].

L'algue verte *Chlamydomonas reinhardtii* est certainement l'organisme chez lequel la photo-réparation de l'ADNpt a été la mieux établie. En effet, la réparation des dimères de pyrimidines est beaucoup plus efficace dans le chloroplaste de *Chlamydomonas* en présence

de lumière qu'à la noirceur [167]. Chez cette algue, le gène *PHR2* encode une CPD photolyase localisée à la fois au chloroplaste et au noyau [168,169]. La photo-réparation des CPDs chloroplastiques a aussi été démontrée chez le maïs, le soya et le riz, ce dernier encodant une photolyase ciblée à la fois au noyau, au chloroplaste et à la mitochondrie [170-172].

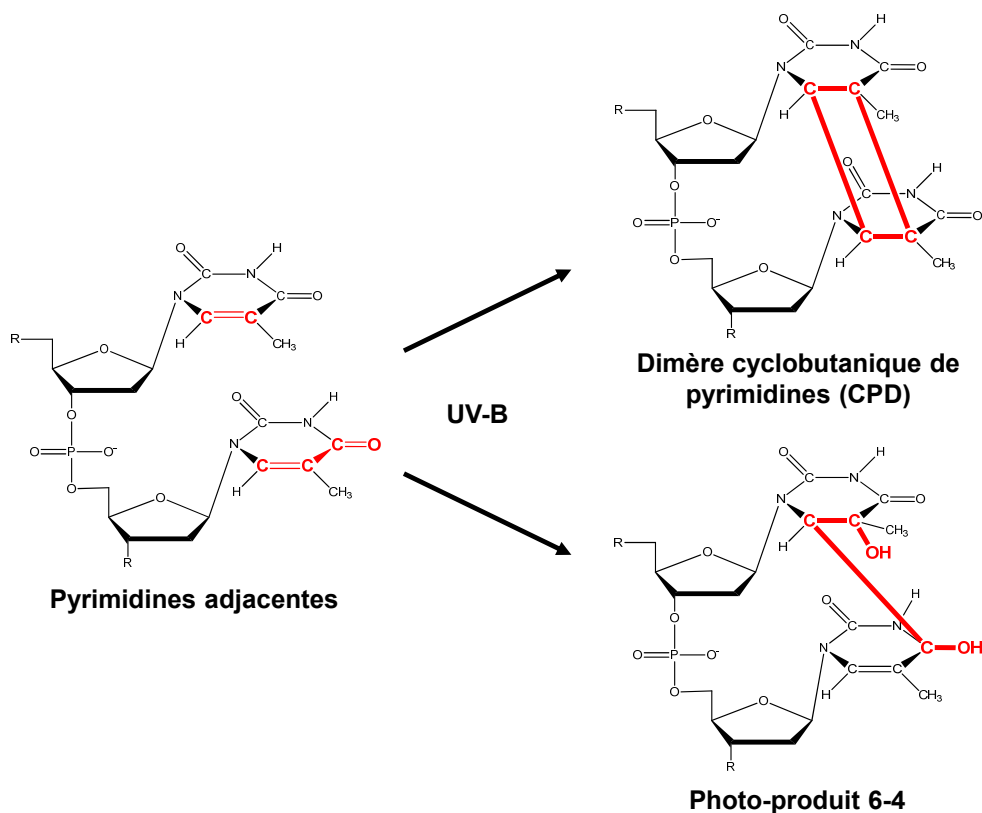


Figure 9. Les rayons UV-B mènent à la formation de liens covalents entre des pyrimidines adjacentes. Ces liens covalents forment des dimères cyclobutaniques de pyrimidines (CPD) et des photo-produits 6-4. Cette figure est adaptée de [74].

L'existence d'un mécanisme de photo-réparation des dimères de pyrimidines dans les chloroplastes d'*Arabidopsis* est cependant beaucoup plus controversée. En effet, les CPDs chloroplastiques ne semblent être que très peu réparés chez de jeunes plants d'*Arabidopsis* de cinq jours, même lorsque les tissus sont exposés à la lumière [173]. Ce résultat est en accord avec une localisation exclusivement nucléaire de la CPD photolyase d'*Arabidopsis*, AtPHR1 [174], et son apparente absence de séquence de ciblage au chloroplaste [175].

Pourtant, des plants de quatorze jours exposés à la lumière bleue réparent tout de même efficacement les CPDs et les photo-produits 6-4 dans leurs chloroplastes [61]. À cet égard, une découverte surprenante a révélé qu'une enzyme de la famille des cryptochromes répare les CPDs chloroplastiques [176,177].

Chez *Arabidopsis*, les cryptochromes CRY1 et CRY2, des photorécepteurs de la lumière bleue, régulent de nombreux processus incluant le photopériodisme, la floraison et la dé-étiolation [revu dans 178]. Curieusement, bien que n'ayant aucune activité photolyase [179,180], les cryptochromes possèdent une séquence primaire, une structure tridimensionnelle et une composition en chromophores similaires aux photolyases [180-182]. L'étude phylogénique des cryptochromes suggère d'ailleurs que ceux-ci partagent un ancêtre commun avec les photolyases [182].

Chez *Arabidopsis*, un troisième cryptochrome, CRY3, a été récemment identifié. Au contraire des autres cryptochromes, CRY3 lie de façon non-spécifique l'ADN du plastide et de la mitochondrie [176,183]. Même si aucune activité photolyase n'a été initialement observée pour ce cryptochrome [176], il a ensuite été démontré que CRY3 possède la capacité de réparer les CPDs, mais uniquement au sein d'une séquence d'ADNss [184,185]. En effet, il semble qu'au fil de l'évolution, CRY3 a conservé son activité photolyase, mais a perdu la capacité à faire basculer les CPDs hors de la double-hélice d'ADN [177]. Il apparaît donc probable que CRY3 participe à la stabilité du génome chloroplastique, des séquences d'ADNss étant constamment générées lors des processus de transcription et de réplication [177].

1.6.3 Les mécanismes de réparation des brisures double-brin de l'ADN

Puisque les rayons UV-B et les ROS n'altèrent que localement la séquence de nucléotides, leurs conséquences demeurent bénignes d'un point de vue génomique. Contrairement à ces types de dommages, les DSBs menacent l'intégrité génomique par la rupture de la double-hélice d'ADN et peuvent induire de graves réarrangements chromosomiques. Ces lésions possèdent donc un fort potentiel déstabilisateur d'un point de vue génomique. Les DSBs

peuvent être générés par l'exposition à des radiations ionisantes, à des agents chimiques endommageant l'ADN ou par des sources endogènes, telles les sites fragiles de l'ADN et les nucléases [revu dans 186]. Afin d'éviter les conséquences désastreuses pouvant être associées aux DSBs, plusieurs processus cellulaires surveillent l'apparition des DSBs et assurent leur réparation. Ceux-ci incluent les mécanismes de recombinaison homologue et les mécanismes reliant deux extrémités d'ADN. D'autre part, un type particulier de DSB est généré lors de l'effondrement de fourches de réplication. Ces DSBs ne possèdent qu'une seule extrémité et nécessitent des mécanismes de réparation particuliers. La prochaine section porte sur les principaux mécanismes réparant les DSBs, sur les évidences de leur existence au plastide et sur les principales protéines du plastide impliquées dans ces voies de réparation.

1.6.3.1 La réparation des DSBs : les mécanismes de recombinaison homologue

Les mécanismes réparant les DSBs générés lors de la fracture de l'hélice d'ADN peuvent être divisés en deux principales catégories : les mécanismes de recombinaison homologue et les mécanismes de religation des deux extrémités du bris. À la différence des mécanismes de religation, les mécanismes de recombinaison homologue re-synthétisent la séquence nucléotidique manquante en utilisant un second duplex d'ADN homologue comme matrice. Ces mécanismes de recombinaison homologue sont ainsi qualifiés de « conservateurs », ceux-ci restituant habituellement la séquence d'ADN d'origine. Néanmoins, l'utilisation d'un duplex homologue comme matrice peut mener à un phénomène nommé « perte d'hétérozygotie », une partie de la séquence d'ADN de la molécule endommagée pouvant être convertie en la séquence du duplex homologue.

Les mécanismes de recombinaison homologue sont initiés par la dégradation des brins 5' de chacune des extrémités du bris. Cette dégradation, également appelée résection, génère des extrémités 3' libres simple-brin très recombino-géniques permettant la synthèse d'ADN. Ces extrémités 3' simple-brin sont ensuite recouvertes par des protéines de liaison à l'ADNss (SSB, *Single-strand DNA-binding protein*, chez les procaryotes; RPA, *Replication protein A*, chez les eucaryotes) qui confèrent une structure linéaire à l'ADNss et préviennent son

appariement non-spécifique. Les protéines de liaison à l'ADNss sont subséquemment remplacées par des recombinaases (RecA chez les procaryotes; RAD51 chez les eucaryotes) assurant l'invasion d'un duplex d'ADN homologue par la ou les extrémités 3' du bris. Une fois le duplex homologue envahi et une boucle D formée, une ADN polymérase synthétise les séquences nucléotidiques manquantes. Enfin, la séparation des duplex, ou la résolution des jonctions de Holliday, restitue deux duplex d'ADN intacts [revu dans 187].

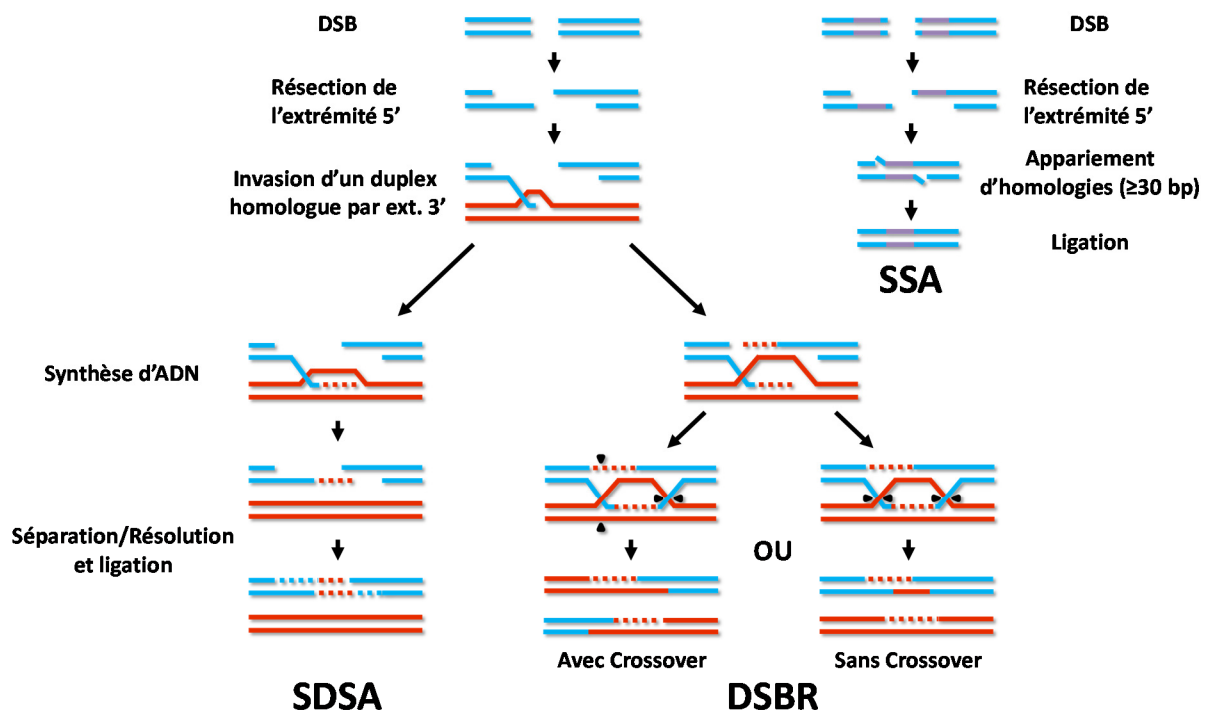


Figure 10. La réparation d'un DSB par recombinaison homologue. Suite à la résection des extrémités 5', l'invasion d'un duplex d'ADN homologue (lignes rouges) permet l'établissement d'une boucle D. La réparation du DSB peut ensuite être complétée par le mécanisme de SDSA (*Synthesis-Dependent Strand Annealing*) ou de DSBR (*Double-Strand Break Repair*). Le mécanisme de SSA (*Single-Strand Annealing*) permet également la réparation d'un DSB. Les lignes violettes représentent des séquences répétées d'au moins 30 nucléotides. Les pointillés représentent des sections d'ADN nouvellement synthétisées. Cette figure est adaptée de [74].

Trois principaux mécanismes de recombinaison homologue ont la capacité de réparer les DSBs : le SDSA (*Synthesis-Dependent Strand Annealing*), le DSBR (*Double-Strand Break Repair*)

et le SSA (*Single-Strand Annealing*). Le SDSA se caractérise par l'envahissement d'un duplex d'ADN homologue par une seule des extrémités 3' du DSB, puis par la synthèse de la séquence d'ADN chevauchant le site du bris (Fig. 10). La dissociation de ce complexe et le ré-appariement du brin nouvellement synthétisé à la seconde extrémité du bris permet finalement le remplissage des régions d'ADNss et la complétion de la réparation. Tout comme le SDSA, le DSBR débute par l'envahissement d'un duplex d'ADN homologue par une des extrémités 3' du DSB. Toutefois, à la différence du SDSA, une synthèse d'ADN est également initiée à partir de la seconde extrémité 3' du DSB, celle-ci utilisant comme matrice le brin déplacé du duplex homologue (Fig. 10). Puisque le mécanisme de DSBR mène à la formation de deux jonctions de Holliday, la restitution de deux duplex d'ADN intacts dépend de la résolution de chacune de ces jonctions. Selon l'orientation de ces résolutions, le mécanisme de DSBR peut résulter en la formation d'un transfert réciproque d'ADN (*crossover*), dans lequel les séquences chromosomiques en aval du bris sont interchangeées. Quant au processus de SSA, celui-ci nécessite, suite à la résection des brins, l'appariement des deux extrémités 3' libres simple-brin par le biais de séquences homologues d'au moins 30 nucléotides (Fig. 10) [188]. Le DSB est ensuite réparé par le clivage des sections d'ADN non-appariées et par la religation des brins. Cependant, même si le SSA est généralement classé parmi les mécanismes de recombinaison homologue, ce processus ne peut être considéré comme un mécanisme de réparation des DSBs conservant la séquence initiale de l'ADN. En effet, la fusion de ces deux séquences répétées génère habituellement des délétions de part et d'autre du DSB [revu dans 188].

La réparation des DSBs par recombinaison homologue constitue sans aucun doute le mécanisme de réparation de l'ADN le plus étudié dans le chloroplaste. En effet, les premières évidences ayant suggéré l'existence de recombinaison homologue dans le chloroplaste ont été obtenues il y maintenant plus de 30 ans. Il avait été conclu que des événements de recombinaison entre les deux IRs étaient à l'origine des deux isomères génomiques composant l'ADNpt [34] (voir section 1.3.2). Des événements de recombinaison entre les IRs ont aussi été observés chez l'algue verte *Chlamydomonas reinhardtii* [189]. Chez cette algue,

le gène encodant l'ADN ribosomal 23S est localisé au sein des IRs et contient un intron du groupe I. Cet intron contient, à son tour, un cadre ouvert de lecture encodant l'endonucléase I-Crel. Cette endonucléase clive de manière très spécifique une séquence nucléotidique de 22 paires de bases située à la jonction exon-exon du gène 23S (c'est-à-dire une copie du gène 23S ne possédant pas l'intron du groupe I). Ainsi, il a été observé que des ADNc du gène 23S intègrent de manière systématique l'intron du groupe I lorsque transformés dans le chloroplaste de *Chlamydomonas* [189]. Il en a donc été conclu que suite au clivage de l'ADNc par I-Crel, le DSB généré est réparé en utilisant comme matrice une copie génomique du gène 23S [190]. De tels résultats démontrent donc l'existence du processus de recombinaison homologue dans le chloroplaste. Les trois principaux mécanismes de recombinaison homologue (DSBR, SDSA et SSA) réparent d'ailleurs les DSBs dans les plastides de *Chlamydomonas* [191].

1.6.3.2 La réparation des DSBs : les mécanismes de religation

La seconde catégorie de mécanismes réparant les DSBs procède par la religation des deux extrémités du bris. Le mécanisme de NHEJ (*Non-Homologous End Joining*) constitue le principal mécanisme de cette catégorie (Fig. 11). Tel que son nom l'indique, le NHEJ répare les DSBs par la ligation de deux extrémités d'ADN sans égard pour leur homologie et, conséquemment, peut mener à la formation de réarrangements chromosomiques [revu dans 187]. De très courtes microhomologies d'un à quatre nucléotides facilitent néanmoins l'amarrage et la religation des deux extrémités du bris [192,193]. Même si le NHEJ est sujet aux erreurs, ce mécanisme parvient tout de même à réapparier les bonnes extrémités du bris dans la majorité des cas. En effet, de façon à éviter les réarrangements génomiques, le complexe protéique de réparation tente de retenir à proximité les deux extrémités d'un même bris [revu dans 186]. À ce jour, aucune protéine responsable de la réparation des DSBs par NHEJ n'a été identifiée au chloroplaste. À priori, ceci semble en accord avec une apparente incapacité des chloroplastes du tabac et de *Chlamydomonas* à réparer un DSB par NHEJ [191,194]. Cette conclusion a toutefois été contredite par l'observation de réarrangements génomiques chloroplastiques générés en absence d'homologie chez

Arabidopsis [27,28]. Il apparaît donc probable qu'un mécanisme actif de NHEJ répare les DSBs dans ce compartiment cellulaire.

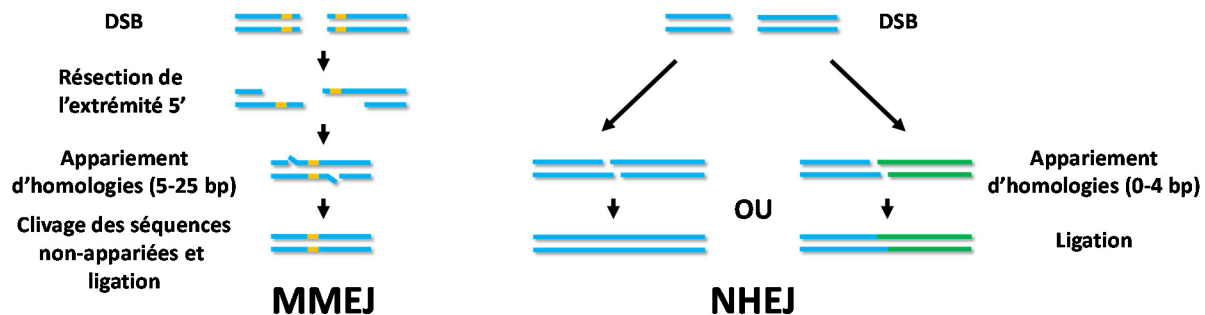


Figure 11. La réparation d'un DSB par religation des deux extrémités du bris.

Lors du processus de MMEJ, la résection des extrémités 5' du bris libère deux extrémités 3' libres simple-brin. Les deux extrémités 3' libres simple-brin se ré-apparient ensuite grâce à des microhomologies de 5 à 25 nucléotides (traits jaunes). Le clivage des séquences non-appariées et la religation des brins finalisent le processus de réparation. Le MMEJ mène habituellement à la délétion de la séquence d'ADN située entre les deux microhomologies. Lors du processus de NHEJ, de très courtes microhomologies d'au plus quatre nucléotides facilitent le ré-appariement des deux extrémités du bris. Les brins sont ensuite reliés de manière à réparer le bris. Puisque le NHEJ relie les deux extrémités du bris sans égard pour leur homologie, ce processus mène parfois à des translocations (ligation avec un duplex d'ADN non-apparié, duplex d'ADN vert dans la figure).

LE NHEJ n'est pas le seul mécanisme reliant les extrémités d'un DSB. Chez la levure et les mammifères, l'inactivation de composantes essentielles du processus de NHEJ n'abolit pas entièrement l'activité de ligation des extrémités d'ADN [revu dans 188]. Toutefois, dans de telles conditions, des microhomologies de 5 à 25 nucléotides sont généralement observées à la jonction de la nouvelle molécule d'ADN reliée. Pour cette raison, ce mécanisme de réparation des DSBs a été nommé *Microhomology-Mediated End Joining* (MMEJ). Par ailleurs, certaines évidences suggèrent que le MMEJ serait actif même lorsque la machinerie de NHEJ est parfaitement fonctionnelle [revu dans 195]. Le MMEJ fonctionne de manière très similaire au processus de SSA, mais utilise, contrairement à celui-ci, des microhomologies afin d'apparier les deux extrémités du bris (Fig. 11) [revu dans 188]. De plus, tout comme le SSA,

le MMEJ est associé à des délétions et à de complexes réarrangements génomiques. Ces deux mécanismes sont pourtant indépendants, ceux-ci ne nécessitant pas la même machinerie protéique [196]. Dans les chloroplastes d'*Arabidopsis*, la réparation d'un DSB induit par une endonucléase génère de nombreuses délétions dépendantes de microhomologies de part et d'autre du bris, suggérant que le mécanisme de MMEJ en serait à l'origine [28].

1.6.3.3 La réparation des DSBs à une seule extrémité

En plus des DSBs engendrés par la cassure de la double hélice d'ADN, un second type de bris, nommé DSB à une seule extrémité (*One-Ended DSB*), menace l'intégrité génomique. Ce type de bris est principalement généré lors de l'effondrement de fourches de réplication. De tels effondrements sont particulièrement fréquents en conditions de stress réplcatif, c'est-à-dire lorsque la progression des ADN polymérases de réplication est ralentie ou inhibée. En de telles conditions, les fourches de réplication bloquées peuvent être converties en DSB à une seule extrémité, dans une ultime tentative de redémarrer la réplication par recombinaison [revu dans 152,197]. Outre le stress réplcatif, des DSBs à une seule extrémité sont également produits lors de la rencontre entre une fourche de réplication et des bris simple-brin de l'ADN (Fig. 12).

Le réplication induite par un bris (*Break-Induced Replication*, BIR, également nommé réplication dépendante de la recombinaison) assure la réparation de ce type de DSB par le redémarrage de la fourche de réplication. De manière similaire aux mécanismes de recombinaison homologue, le BIR est initié par la résection de l'extrémité 5' du bris (Fig. 12). Une recombinase (RecA chez les procaryotes; RAD51 chez les eucaryotes) permet ensuite l'invasion d'un duplex d'ADN homologue par l'extrémité 3' et la formation d'une boucle D. La suite du processus diffère toutefois de la recombinaison homologue. En effet, une fourche de réplication complète est ré-établie au site de la boucle D et synthétise l'ADN autant sur le brin continu que discontinu [198,199]. Cette fourche de réplication nouvellement établie réplique alors l'ADN jusqu'à l'extrémité du chromosome afin de compléter la ronde de réplication interrompue.

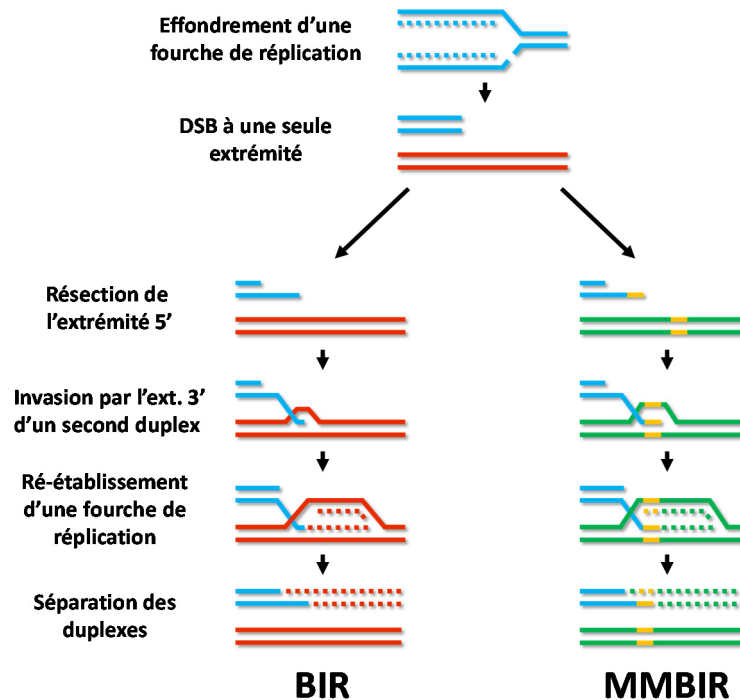


Figure 12. La réparation d'un DSB à une seule extrémité par les processus de BIR et de MMBIR. L'effondrement de fourches de réplication et la rencontre d'un bris simple-brin dans l'ADN par une fourche de réplication mènent à la formation de DSB à une seule extrémité. Lors du BIR, la résection de l'extrémité 5' du bris et l'invasion d'un duplex d'ADN homologue (lignes rouges) permettent l'établissement d'une fourche de réplication. La progression de la fourche de réplication jusqu'à l'extrémité du chromosome assure la complétion de la ronde de réplication. Lors du MMBIR, l'invasion d'un second duplex d'ADN est guidée par une microhomologie (tirets jaunes). Cette dépendance du MMBIR envers de courtes microhomologies mène à l'établissement de fourches de réplication à l'endroit de duplex d'ADN non-homologues (duplex d'ADN vert dans la figure). Les pointillés représentent des sections d'ADN nouvellement synthétisées.

Dans certains cas, le BIR peut être initié en absence d'une recombinaise. Le site d'invasion d'un second duplex d'ADN est alors guidé par une courte microhomologie. Par conséquent, ce processus se trouve à être beaucoup moins fidèle que le BIR et génère de complexes réarrangements génomiques, tels que des duplications, des délétions, des inversions ou des translocations [revu dans 200,201] (Fig. 12). Afin de souligner cette dépendance envers des séquences de microhomologies, ce type particulier de BIR est nommé MMBIR (*Microhomology-Mediated Break-induced Replication*) [200]. De tels événements sont

observés dans le génome du plastide des végétaux et sont particulièrement fréquents lors d'un dysfonctionnement de la machinerie de réparation de l'ADNpt [27,43] (voir section 1.6.3.4). Des plants d'*Arabidopsis* qui ont poussés en présence de ciprofloxacine, un inhibiteur des gyrases introduisant des DSBs dans l'ADN des organites, accumulent d'ailleurs des réarrangements génomiques induits par des microhomologies dans leur génome mitochondrial et chloroplastique [29]. La séquence de microhomologie observée à la jonction de ces réarrangements ne permet cependant pas de discriminer sans équivoque les mécanismes de MMBIR et de MMEJ. Néanmoins, puisque le mécanisme de MMEJ est essentiellement associé à des délétions, l'observation fréquente de duplications dans le génome du plastide renforce l'idée qu'un mécanisme de MMBIR soit actif dans ce compartiment.

1.6.3.4 Les protéines assurant la réparation des DSBs dans le chloroplaste

A) Les homologues de RecA

Chez les procaryotes, les trois principaux mécanismes de recombinaison homologue (SDSA, DSBR, BIR) s'amorcent par la liaison des protéines RecA aux extrémités 3' simple-brin du bris. Cette enzyme catalyse l'échange de brins entre une extrémité 3' d'ADNss et un duplex d'ADN homologue de façon à former une boucle D (Fig. 10 et 12). Des homologues de RecA sont aussi retrouvés dans les chloroplastes et les mitochondries de plusieurs végétaux, dont *Arabidopsis*, le pois *Pisum sativum*, la mousse *Physcomitrella patens* et l'algue verte *Chlamydomonas reinhardtii* [202-206]. Chez *Arabidopsis*, trois homologues de RecA sont encodés dans le génome nucléaire. La fusion de leurs séquences de ciblage à la protéine fluorescente verte (GFP) a établi que RECA1 et RECA3 sont dirigés au chloroplaste et à la mitochondrie, respectivement, et que RECA2 est ciblé aux deux organites [207].

La fonction des homologues chloroplastiques de RecA a été amplement étudiée au début des années 1990. Il a entre autres été observé que chez *Arabidopsis*, une protéine chloroplastique apparentée à la protéine RecA bactérienne est surexprimée lors d'un traitement avec des agents endommageant l'ADN [44,202]. Chez le pois, cette protéine

possède une activité d'échange de brins d'ADN strictement dépendante d'un duplex d'ADN homologue [208]. L'importance du processus de recombinaison homologue pour la réparation des dommages à l'ADN du plastide a également été démontrée par l'emploi, chez *Chlamydomonas*, d'une lignée *reca* mutante hypersensible aux dommages à l'ADN [209]. Plus récemment, l'étude d'un mutant du gène *RECA1* d'*Arabidopsis* a montré que cette mutation est associée à une altération de la structure du génome du plastide et à une augmentation des niveaux d'ADNss dans ce compartiment [210]. Ce mutant présente aussi une hypersensibilité à la ciprofloxacine [210]. Globalement, ces résultats renforcent l'idée que l'homologue chloroplastique de RecA participe activement au maintien de l'intégrité génomique. Pour le moment, aucune étude n'a toutefois vérifié si la mutation *reca1* occasionne de l'instabilité génomique dans le plastide.

Par ailleurs, l'étude des homologues mitochondriaux de RecA supporte l'importance de cette famille de protéines pour le maintien de l'intégrité génomique dans les organites. Comparativement au génome du plastide, de très nombreuses séquences répétées de longueur intermédiaire (≥ 50 bases) sont retrouvées dans l'ADN mitochondrial et peuvent mener à des événements de recombinaison illégitime [revu dans 211]. Chez *Arabidopsis* et *Physcomitrella patens*, les mutations d'*AtRECA3* et de *PpRECA1* provoquent toutes deux l'accumulation de réarrangements génomiques mitochondriaux causés par des événements de recombinaison illégitime entre ce type de répétitions [207,212,213]. Le mutant *atreca3* présente d'ailleurs une réduction de son taux de réparation des DSBs et est, par conséquent, hypersensible aux agents endommageant l'ADN tels que la ciprofloxacine [213]. Chez *Arabidopsis*, aucun phénotype particulier n'est pourtant associé à la mutation *reca3*. Étonnamment, la mutation du gène *RECA2*, ciblé à la fois à la mitochondrie et au chloroplaste, empêche la germination des plants d'*Arabidopsis* [207,213]. Ce mutant parvient néanmoins à germer en présence d'un apport externe de sucre, mais sa croissance est presque immédiatement interrompue au stade des cotylédons, c'est-à-dire les deux premières feuilles issues de la graine. Même si cette protéine est ciblée dans les deux organites, cet arrêt de la croissance semble davantage lié au rôle mitochondrial que

chloroplastique de RECA2. En effet, puisque les jeunes plants *reca2* qui poussent en présence de sucre présentent une coloration verte foncée, il semble probable que la biogénèse des chloroplastes ne soit que très peu affectée par cette mutation [213]. De manière générale, les homologues de RecA retrouvés dans les organites d'*Arabidopsis* semblent donc être des acteurs clés du maintien de la stabilité génomique chloroplastique et mitochondriale.

B) Les protéines Whirly

Comme discuté dans les sections précédentes, plusieurs des mécanismes qui assurent la réparation des DSBs requièrent la formation d'une extrémité 3' libre d'ADNss très recombino-génique et sujet à des appariements non-spécifiques. Les extrémités 3' d'ADNss doivent donc être stabilisées de façon à limiter les mécanismes de réparation sujets aux erreurs et favoriser une réparation adéquate. Par ailleurs, de nombreux processus du métabolisme de l'ADN, tels que la transcription et la réplication, nécessitent la formation temporaire d'ADNss. Bien que cet ADNss ne constitue pas un intermédiaire de la réparation des DSBs, il a été proposé que celui-ci favorise les événements de recombinaison illégitime en facilitant l'envahissement des extrémités 3' d'ADNss [29]. Pour ces raisons, l'ADNss est rarement libre dans la cellule et se retrouve en complexe avec des protéines stabilisant l'ADNss. Chez les procaryotes, l'ADNss se retrouve associé aux protéines SSBs (*Single-Strand DNA Binding*). Ces protéines préviennent les événements de recombinaison illégitime [214,215] et orchestrent le processus de réparation des DSBs par recombinaison homologue [revu dans 216].

Dans les organites des végétaux, cette fonction de stabilisation de l'ADNss est accomplie, entre autres, par les protéines de la famille Whirly. Trois protéines Whirly sont encodées dans le génome nucléaire d'*Arabidopsis*. Alors que WHY2 est ciblée à la mitochondrie, les protéines WHY1 et WHY3 sont, quant à elles, dirigées au plastide [217]. Ces protéines s'associent sous forme d'homotétramères et lient l'ADNss sans égard pour sa séquence [29]. De plus, lors de leur liaison à de longues molécules d'ADNss, les tétramères de protéines

Whirly exercent de la coopérativité et s'associent en mégasstructures hautement ordonnées [218].

L'importance des protéines WHY1 et WHY3 dans le maintien de la stabilité du génome chloroplastique a été récemment élucidée par une approche de génétique inverse [27]. Chez *Arabidopsis*, la mutation des deux gènes encodant les protéines Whirly du plastide, *why1why3*, provoque l'apparition d'un phénotype de variégation blanche (c'est-à-dire l'apparition de taches blanches dépourvues de chlorophylle sur les feuilles) chez environ cinq pourcents de la population mutante [27]. L'observation des tissus variégés de *why1why3* par microscopie électronique démontre que ce phénotype est associé à une altération de la biogénèse des chloroplastes, les membranes thylakoïdes de ceux-ci ne se développant que très peu [27]. D'autre part, des croisements réciproques ont démontré que ce phénotype se transmet de façon maternelle, suggérant qu'une altération dans le génome du plastide en est responsable [27].

L'analyse du génome chloroplastique a effectivement révélé que le phénotype de variégation observé chez *why1why3* est associé à de l'instabilité génomique chloroplastique, certaines régions du génome étant hautement amplifiées chez les individus variégés [27]. Il a été observé que de très courtes microhomologies d'une quinzaine de nucléotides bordent les régions réarrangées, suggérant que des événements de recombinaison entre ces très courtes répétitions soient en cause dans cette amplification [27]. Des réarrangements génomiques de l'ADNpt ne sont toutefois pas observés chez les plants *why1why3* variégés. Des duplications et des délétions sont en effet observées dans l'ensemble de la population *why1why3* [27]. Les caractéristiques associées à ces réarrangements suggèrent que le mécanisme de MMBIR serait facilité chez les mutants *why1why3*.

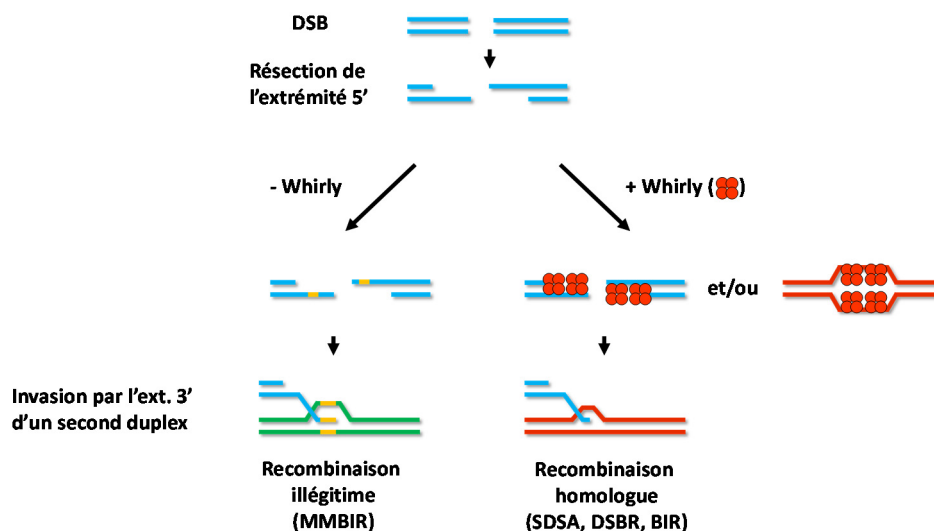


Figure 13. Les protéines Whirly favorisent la réparation adéquate des DSBs en stabilisant l'ADNss et en limitant les appariements non-spécifiques. Lors de la résection des extrémités 5' d'un DSB, des séquences d'ADNss hautement recombinogéniques sont générées. Par leur liaison à l'ADNss, les protéines Whirly limiteraient les appariements non-spécifiques et favoriseraient les voies de recombinaison homologue. En absence de Whirly, de courtes microhomologies seraient suffisantes pour initier des événements de recombinaison illégitime. Les traits rouges et verts représentent, respectivement, un duplex d'ADN homologue et non-homologue. Les courts traits jaunes représentent des microhomologies.

Cette hypothèse est d'ailleurs renforcée par l'observation que les mutants *why1why3* sont hypersensibles à un agent introduisant des DSBs dans l'ADN des organites, la ciprofloxacine, et que ce traitement provoque une accumulation accrue de réarrangements génomiques [29]. Il semble donc qu'en présence de DSBs, les protéines Whirly stabilisent l'ADNss généré lors de la résection des extrémités 5' jusqu'à sa prise en charge par la machinerie de recombinaison homologue (Fig. 13). Il paraît aussi probable que les protéines Whirly stabilisent l'ADNss généré lors des processus du métabolisme de l'ADNpt afin d'éviter les appariements non-spécifiques avec les extrémités 3' d'ADNss. La caractérisation cristallographique de l'interface entre l'ADNss et les protéines Whirly supporte d'autant plus cette hypothèse puisque, dans le complexe nucléoprotéique, les bases azotées ne sont pas accessibles pour effectuer des appariements avec des séquences complémentaires [29].

C) L'ADN polymérase POLIB

Il a été démontré que deux ADN polymérases sont responsables de la réplication de l'ADNpt chez le riz, le tabac et *Arabidopsis* [43,63-66]. Ces deux polymérases n'ont pourtant pas des rôles totalement redondants. En effet, l'expression de *POLIB*, mais pas de *POLIA*, est augmentée chez *Arabidopsis* suite à une exposition au peroxyde d'hydrogène, suggérant une spécialisation de *POLIB* dans la réparation de l'ADN [64]. Cette spécialisation de *POLIB* est d'autant plus appuyée par la caractérisation des mutants *polla* et *pollb* [43]. En plus de démontrer une hypersensibilité à la ciprofloxacine, le mutant *pollb* présente une réduction de son taux de réparation des DSBs dans le chloroplaste et dans la mitochondrie lorsque soumis à cet agent génotoxique [43]. Il a également été démontré que la polymérase *POLIB* et les protéines Whirly chloroplastiques, *WHY1* et *WHY3*, collaborent au maintien de la stabilité génomique du plastide. Une interaction génétique synergique est observée, chez *Arabidopsis*, lors de la combinaison des mutations *why1why3* et *pollb*, l'entièreté de la population triple mutante présentant un important retard de croissance et une pâleur de ses tissus [43]. Cet effet synergique se reflète aussi sur l'accumulation de réarrangements génomiques de l'ADNpt, les plants *why1why3pollb* présentant des niveaux d'instabilité génomique bien supérieurs aux plants *why1why3* [43]. Le génome du triple mutant *why1why3pollb* est effectivement hautement réarrangé par de nombreuses duplications et délétions vraisemblablement associées au mécanisme de MMBIR [43]. Il semble donc probable que la polymérase *POLIB* travaille de concert avec les protéines *WHY1* et *WHY3* à la réparation adéquate des DSBs, de manière à limiter les réarrangements génomiques.

D) Les autres acteurs

Bien que peu documentées, d'autres protéines préviennent l'accumulation de réarrangements génomiques dans le plastide et dans la mitochondrie. Parmi celles-ci, il convient de mentionner la protéine *MSH1* (*MutS homolog 1*), ciblée dans ces deux organites [219,220]. Chez *E. coli*, la protéine *MutS* participe au processus de réparation des mésappariements (*Mismatch repair*). Contrairement à son homologue bactérien, aucune

activité dans le processus de réparation des mésappariements n'a été observée pour MSH1. Il a plutôt été démontré que MSH1 réprime les événements de recombinaison illégitime entre des séquences répétées de longueur intermédiaire (≈ 100 à 500 nucléotides) dans le génome mitochondrial [207,221]. Dans le plastide, MSH1 joue un rôle similaire aux protéines Whirly en réprimant des événements de recombinaison entre de courtes microhomologies [220]. La mutation *msh1* provoque d'ailleurs l'accumulation de réarrangements génomiques mitochondriaux et chloroplastiques en plus d'être associée à un phénotype de variégation blanche [220,222]. Ce phénotype étant transmis de façon maternelle, il semble probable que ces réarrangements génomiques en soient la cause.

Dans la mitochondrie, la protéine OSB1 (*Organellar Single-stranded DNA Binding protein 1*) participe également à la répression de la recombinaison entre des séquences répétées de longueur intermédiaire [223]. Chez *Arabidopsis*, la famille des protéines OSBs comprend également un membre retrouvé au plastide, OSB2. Il est donc possible qu'OSB2 participe également au maintien de la stabilité génomique dans ce compartiment.

D'autre part, il a été récemment démontré qu'un homologue de l'ADN hélicase bactérienne RecG participe au maintien de l'intégrité génomique du plastide et de la mitochondrie chez la mousse *Physcomitrella patens* [224]. Chez *E. coli*, RecG contribue au redémarrage des fourches de réplication bloquées et est, par conséquent, un acteur important au maintien de la stabilité génomique [revu dans 225]. Chez *P. patens*, la mutation *recg* provoque l'accumulation de réarrangements génomiques mitochondriaux et chloroplastiques entre de très courtes microhomologies [224]. Le mécanisme d'action de RECG dans ces compartiments reste toutefois inconnu pour l'instant.

Finalement, un homologue de RAD52, une protéine catalysant la polymérisation de RAD51 sur l'ADNss, a été identifié dans tous les compartiments contenant de l'ADN chez *Arabidopsis* [226]. Cette protéine lie l'ADNss sans aucune spécificité pour sa séquence et, dans la mitochondrie, participe au processus de recombinaison homologue [227]. Sa mutation cause d'ailleurs une diminution du taux de réparation des DSBs par recombinaison homologue et

une augmentation des événements de recombinaison entre de courtes microhomologies [227]. Ainsi, il est tentant de spéculer que cette protéine joue un rôle similaire dans le plastide. Pour le moment, aucune évidence ne suggère toutefois sa participation à la réparation des DSBs dans le plastide.

1.7 Objectifs et approche expérimentale

À mon arrivée au laboratoire, il avait été démontré que les protéines Whirly localisées au plastide, soit *WHY1* et *WHY3*, participent au maintien de l'intégrité génomique de cet organe en réprimant les événements de recombinaison illégitime survenant entre de courtes microhomologies d'ADN. Il avait également été observé que les protéines Whirly travaillent de concert avec l'ADN polymérase *POLIB*, la combinaison des mutations *why1why3pol1b* provoquant une forte accumulation de ce type d'instabilité génomique. Bien que ce mutant éprouve un sévère trouble développemental, les conséquences d'une déstabilisation généralisée de l'ADN par les réarrangements génomiques demeuraient peu comprises. Par ailleurs, outre les protéines Whirly et *POLIB*, il était connu qu'un homologue de la recombinase bactérienne *RecA*, *RECA1*, contribue au maintien de la stabilité génomique chloroplastique. Toutefois, l'interrelation existant entre *RECA1* et les gènes *POLIB*, *WHY1* et *WHY3* demeurait inconnue tant d'un point de vue physiologique que de son importance pour le maintien de la stabilité du génome chloroplastique. D'autre part, une approche PCR avait été utilisée jusqu'alors au laboratoire pour analyser l'instabilité génomique des organites. Une telle approche ne donnait toutefois qu'une image partielle de l'instabilité génomique en plus d'être hautement biaisée et peu quantitative. Dans le cadre de mes études doctorales, nous nous sommes donc donné comme objectifs de :

- I. Identifier les conséquences physiologiques et moléculaires de l'instabilité génomique chloroplastique.
- II. Caractériser les probables relations génétiques existant entre les gènes *RECA1*, *POLIB*, *WHY1* et *WHY3* et leurs implications dans le maintien de l'intégrité génomique du chloroplaste.

- III. Développer une approche quantitative, non biaisée et fournissant un patron global de l'instabilité génomique dans les organites.
- IV. Évaluer, par une telle approche, les principaux types d'instabilité génomique menaçant l'intégrité des génomes des organites.

Afin d'établir les conséquences de l'instabilité génomique chloroplastique, nous avons tiré avantage du mutant *why1why3pollb*, qui possède un génome chloroplastique hautement déstabilisé par des réarrangements génomiques, et de la ciprofloxacine, un composé engendrant un stress génotoxique affectant de manière spécifique les chloroplastes et les mitochondries végétales. Ensuite, nous avons utilisé une approche de génétique inverse afin d'étudier les conséquences de la perte de fonction du gène *RECA1* d'*Arabidopsis*. Nous avons également généré des lignées combinant les mutations *reca1pollb* et *why1why3reca1* dans le but d'évaluer les relations génétiques existant entre ces gènes. Finalement, nous avons tiré profit des technologies de séquençage de nouvelle génération pour développer une approche permettant l'analyse des patrons d'instabilité génomique des organites de manière quantitative et non biaisée.

**CHAPITRE 2. Plastid Genome Instability
Leads to Reactive Oxygen Species
Production and Plastid-to-Nucleus
Retrograde Signaling in *Arabidopsis***

Plastid Genome Instability Leads to Reactive Oxygen Species Production and Plastid-to-Nucleus Retrograde Signaling in *Arabidopsis*

Publié dans Plant Physiology, octobre 2013

Vol. 163, numéro 2, pages 867–881

Étienne Lepage*, Éric Zampini*, et Normand Brisson

*Ces auteurs ont contribué également à l'article

www.plantphysiol.org

Copyright American Society of Plant Biologists.

2.1 Mise en contexte

Bien que les génomes mitochondriaux et chloroplastiques n'encodent que peu de gènes, leur dysfonction entraîne de très sévères phénotypes. Chez l'humain, par exemple, plusieurs centaines de mutations et de réarrangements mitochondriaux ont été associés à de très graves maladies et syndromes [revu dans 228]. Même si les conséquences de ces mutations restent généralement mal comprises, il est fréquemment suggéré que cette instabilité génomique décroît l'efficacité du processus de respiration cellulaire et entraîne un déséquilibre de la chaîne de transport des électrons. En plus, les mutations de l'ADN mitochondrial sont associées, dans de nombreux cas, à une production accrue de ROS et à un déséquilibre rédox cellulaire. Au contraire des mitochondries humaines, les conséquences moléculaires et physiologiques de l'instabilité génomique chloroplastique chez les végétaux sont beaucoup moins comprises. Alors que de graves problèmes développementaux ont été associés à la mutation de certains gènes chloroplastiques [229], les conséquences d'une déstabilisation généralisée de l'ADNpt restaient inconnues. Afin de déterminer les conséquences d'une telle déstabilisation de l'ADNpt, nous avons utilisé la lignée mutante *why1why3pollb* qui présente une accumulation importante de réarrangements génomiques chloroplastiques associés au mécanisme de MMBIR [43].

Dans cet article, j'ai participé aux expériences qui ont permis d'établir un lien entre l'instabilité génomique chloroplastique, le déséquilibre photosynthétique et le stress oxydatif. J'ai aussi participé aux caractérisations physiologiques du mutant *why1why3pollb* et des plantes traitées à la CIP, aux expériences d'amplification PCR des réarrangements génomiques et aux expériences de microscopie présentées dans l'article. De plus, j'ai participé à toutes les étapes de l'analyse transcriptomique présentée dans cet article. É. Lepage a contribué aux expériences de mesure des paramètres physiologiques, de microscopie, d'amplification PCR des réarrangements génomiques et de séquençage de l'ARN. Finalement, l'interprétation des résultats, l'écriture et la révision du manuscrit ont été effectuées conjointement par É. Lepage, N. Brisson et moi-même.

2.2 Abstract

The plastid genome is highly conserved among plant species, suggesting that alterations of its structure would have dramatic impacts on plant fitness. Nevertheless, little is known about the direct consequences of plastid genome instability. Recently, it was reported that the plastid Whirly proteins, WHY1 and WHY3, and a specialized type-I polymerase, POLIB, act as safeguards against plastid genome instability in *Arabidopsis thaliana*. In the present study, we use ciprofloxacin, an organelle double-strand break inducing agent, and the *why1why3polib-1* variegated mutant to evaluate the impact of generalized plastid DNA instability. First, we show that in *why1why3polib-1* and ciprofloxacin treated plants, plastid genome instability is associated to increased ROS production. Then, using different light regimens, we show that the elevated ROS production correlates with the appearance of a yellow-variegated phenotype in the *why1why3polib-1* population. This redox imbalance also correlates to modifications of nuclear gene expression patterns, which in turn leads to acclimation to high light. Taken together, these results indicate that plastid genome instability induces an oxidative burst that favors, through nuclear genetic reprogramming, adaptation to subsequent oxidative stresses.

2.3 Introduction

As opposed to most non-photosynthetic eukaryotes, plants possess an additional endosymbiotic organelle, the plastid, which is the seat of photosynthesis. In addition to this essential process, plastids have a variety of other important biological roles. Of these, chloroplasts serve as sensors of environmental conditions and guide the plant toward proper adaptation [230]. Indeed, as a result of its light-harvesting capacity, the chloroplast is directly associated with light-absorption and, accordingly, is ideally placed to detect environmental fluctuations [82]. The fluctuating environmental conditions are perceived through several important molecular checkpoints, such as the redox state of the photosynthetic electron transport chain (PET), often measured as the excitation pressure of the photosystem II (PSII), or the concentration of reactive oxygen species (ROS) produced by the PET [231].

Even though photosynthesis serves as the main energy source for autotrophic organisms, this process also constantly generates ROS by-products that threaten all cellular components by their strong propensity for oxidation [78]. ROS production is notably exacerbated by abiotic stresses, such as a sudden exposure to high light intensities, which leads to excessive photon absorption, over-reduction of the plastoquinone (PQ) pool, and electron transfer to molecular oxygen. If improperly scavenged, ROS can have a variety of deleterious effects such as base oxidation, lipid peroxidation, and in some cases cell death [232,233].

Beyond their deleterious effects, ROS also act as important signaling molecules, especially under stress conditions. Indeed, ROS-dependent signaling has been observed in the vast majority of biotic and abiotic stress responses [234-236]. Recently, different studies have highlighted specific reprogramming of the nuclear transcriptome depending on the site of production within the cell and the type of ROS involved [100,101,237]. When produced in the chloroplast, ROS initiate plastid-to-nucleus retrograde signaling and modify nuclear and chloroplast gene expression patterns to adjust the photosynthetic activity. For instance, transcriptome alterations during chloroplast oxidative stress often include repression of the light-harvesting complexes and accumulation of high levels of antioxidant proteins [85,89].

The negative correlation between PET efficiency and chloroplast ROS production suggests that any disequilibrium in PET homeostasis can lead to redox imbalance. Accordingly, mutations of nuclear genes involved in the quality control of the PET proteins, such as the FtsHs and Deg proteases, induce developmental defects and redox imbalance [81,238-240]. The chloroplast gene expression status acts as an additional checkpoint to ensure PET homeostasis. Indeed, although plastid DNA (ptDNA) only encodes a small fraction of chloroplast proteins, the retained plastid genes are mostly involved in photosynthesis [23]. Thus, inhibition of chloroplast gene expression induced by mutations or chemical treatments leads to impairment of photosynthesis and increased chloroplast ROS production [127,241,242]. This suggests that the maintenance of plastid genome stability is crucial to ensure PET efficiency. However, although the function of many nuclear genes involved in

chloroplast DNA repair is emerging, little is known about the direct outcome of generalized plastid genome instability on chloroplast function.

Among the different types of DNA damages leading to genome instability, double-strand breaks (DSBs) are considered the most threatening since they can give rise to severe chromosomal rearrangements if improperly repaired [191]. This type of damage can spontaneously arise in the chloroplast genome as a consequence of replication fork collapse or can be specifically induced by ciprofloxacin (CIP), a gyrase inhibitor that produces DSBs in a replication-dependent manner [43,243]. In *Arabidopsis*, PolIA and PolIB, two type-I DNA polymerases, are proposed to perform chloroplast DNA replication [43,64]. The mutation of either of the genes encoding these proteins is enough to cause replication stress and PolIB has been demonstrated to have additional roles in DNA repair [43] and mitochondrial homeostasis [244]. In addition to PolIB, plant organelles possess Whirly proteins that act as safeguards of plastid genome stability by binding and stabilising single-stranded DNA, thereby limiting the occurrence of illegitimate recombination [27,29]. *Arabidopsis* mutants for the two plastid Whirly genes, *WHY1* and *WHY3*, are sensitive to CIP and exhibit ptDNA rearrangements that are thought to occur through a microhomology-mediated break-induced replication (MMBIR) mechanism [27,29,200]. Interestingly, in the *why1why3* double mutant, the absence of chloroplast Whirly proteins causes variegation in about 5 % of the plant population [27]. Moreover, the mutation of *PolIB* has, together with the mutations of *WHY1* and *WHY3*, a synergistic effect on the appearance of variegation and accumulation of ptDNA rearrangements. The *why1why3polib-1* triple mutant displays a 100 % penetrance yellow-variegated phenotype and accumulates even larger amounts of MMBIR-associated rearrangements than *why1why3* plants [43].

Here we report that ptDNA instability, induced by two distinct approaches, leads to an altered cellular redox status. We show that CIP treatment, which induces DSBs in the chloroplast, and the *why1why3polib-1* mutations lead to decreased PET efficiency and elevated chloroplast ROS production in *Arabidopsis*. Our results also reveal that ptDNA rearrangements and increased ROS levels in *why1why3polib-1* trigger chloroplast-to-nucleus

retrograde signaling and overexpression of many abiotic stress-related genes. Finally, we show that this ROS-dependent genetic reprogramming promotes an acquired adaptation to high light.

2.4 Results

2.4.1 Plastid Genome Instability Negatively Impacts Photosynthetic Efficiency

Leaf variegation has been observed in many different mutants of *Arabidopsis*. The severity of this phenotype varies among these mutants and is sometimes associated to dwarfism [27,43,220,245,246]. In *Arabidopsis*, the variegation observed in *why1why3* and *why1why3pollb-1* (Fig. 14A-B) has been linked to DNA rearrangements produced by the MMBIR pathway in the chloroplast genome (Fig. 23, page 85) [27,43]. While the *why1why3* plants present a moderate amount of MMBIR-associated rearrangements, these rearrangements are present at much higher levels in the triple mutant *why1why3pollb-1* (Fig. 14C) [43]. Interestingly, the high levels of rearrangements in *why1why3pollb-1* do not exacerbate the *why1why3* white variegation phenotype. Instead, a distinct yellow-variegated phenotype is observed in *why1why3pollb-1* (Fig. 14B and Fig. 24, page 86). The fact that *why1why3pollb-1* also presents white variegation in an equivalent fraction of its population further supports that white and yellow variegation are distinct phenotypes (Fig. 25, page 87).

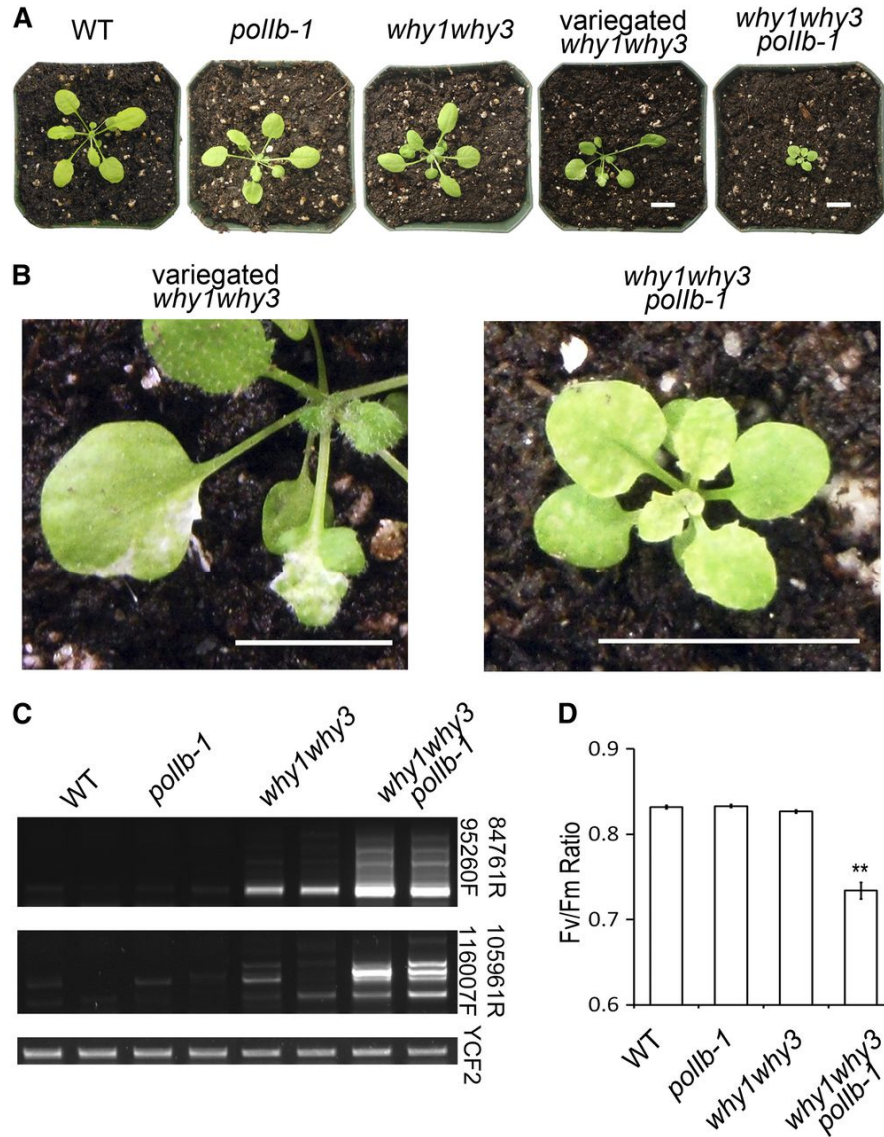


Figure 14. Plastid DNA Rearrangements in *why1why3pollb-1* Lead to Photosynthetic Impairment and Yellow Variegation. (A) Representative photographs of 21-day-old wild type and *pollb-1*, *why1why3* and *why1why3pollb-1* *Arabidopsis* mutants grown under normal light conditions. White bars represent 1 cm. (B) Enlargements showing the variegated sectors of *why1why3* and *why1why3pollb-1*. White bars represent 1 cm. (C) Representative outward-facing PCR reactions showing the abundance of MMBIR-associated plastid DNA rearrangements in 14-day-old seedlings of the mutant lines presented in (A). PCR reactions were carried out using the indicated primer pairs. The amount of DNA loaded in each reaction was equilibrated relative to the amplification of a *YCF2* fragment. Every reaction was performed on three different biological samples and a representative biological duplicate is presented. (D) Fv/Fm ratios for 21-day-old plants of the indicated genotype

grown under normal light conditions. The error bars represent the standard error of the mean (SEM). Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type. Every measurement was performed independently on at least five biological samples.

Although the roles of Why1, Why3 and PolIB in plastid genome stability are becoming well established, the consequences of the ptDNA rearrangements associated with their mutations are still poorly understood. Since the plastid genome encodes many components of the PET, it is conceivable that plastid genome instability could impair photosynthesis. Therefore, to evaluate the potential impacts of MMBIR-associated ptDNA rearrangements on the PET photosynthetic capacity, the efficiency of the PSII was measured in wild type, *pollb-1*, *why1why3* and *why1why3pollb-1* plants grown for four weeks under normal light conditions. Measurement of the Fv/Fm ratio, a widely used parameter for the measurement of the PSII efficiency, indicates that the *why1why3pollb-1* mutant exhibits a significant decrease in its photosynthetic capacity (Fig. 14D). To confirm the link between ptDNA rearrangements and photosynthetic defects, DSBs were introduced in the organelle genomes of WT plants by treatment with ciprofloxacin (CIP). Similarly to *why1why3pollb-1*, wild type plants treated with CIP presented a significant decrease in their PSII efficiency as revealed by a lower Fv/Fm ratio (Fig. 26, page 88). No such decrease was observed when the plants were treated with novobiocin (NOVO) (Fig. 26), a gyrase inhibitor that induces neither DSBs nor ptDNA rearrangements [29,247]. Taken together, the lowered Fv/Fm ratios observed in *why1why3pollb-1* and upon CIP treatment clearly suggest that ptDNA rearrangements negatively impact PET performance.

2.4.2 Plastid Genome Instability Leads to Generation of ROS in the Chloroplast

Since PET defects and imbalance are often associated with increased ROS production [81,238], we verified whether the ptDNA rearrangements observed in Figure 14C are also associated with oxidative stress in the chloroplasts. Leaves from four-week-old *Arabidopsis* plants were incubated in the presence of the general oxidative stress fluorescent probe H₂DCFDA. Microscopic detection of the oxidized fluorescent form of the molecule revealed

ROS-generating foci in *why1why3pollb-1* plants, but not in wild type, *why1why3* and *pollb-1* plants (Fig. 15A). Moreover, the fluorescent H₂DCFDA colocalizes with chlorophyll and structures resembling plastids, suggesting that chloroplasts are the origin of this redox imbalance (Fig. 15B).

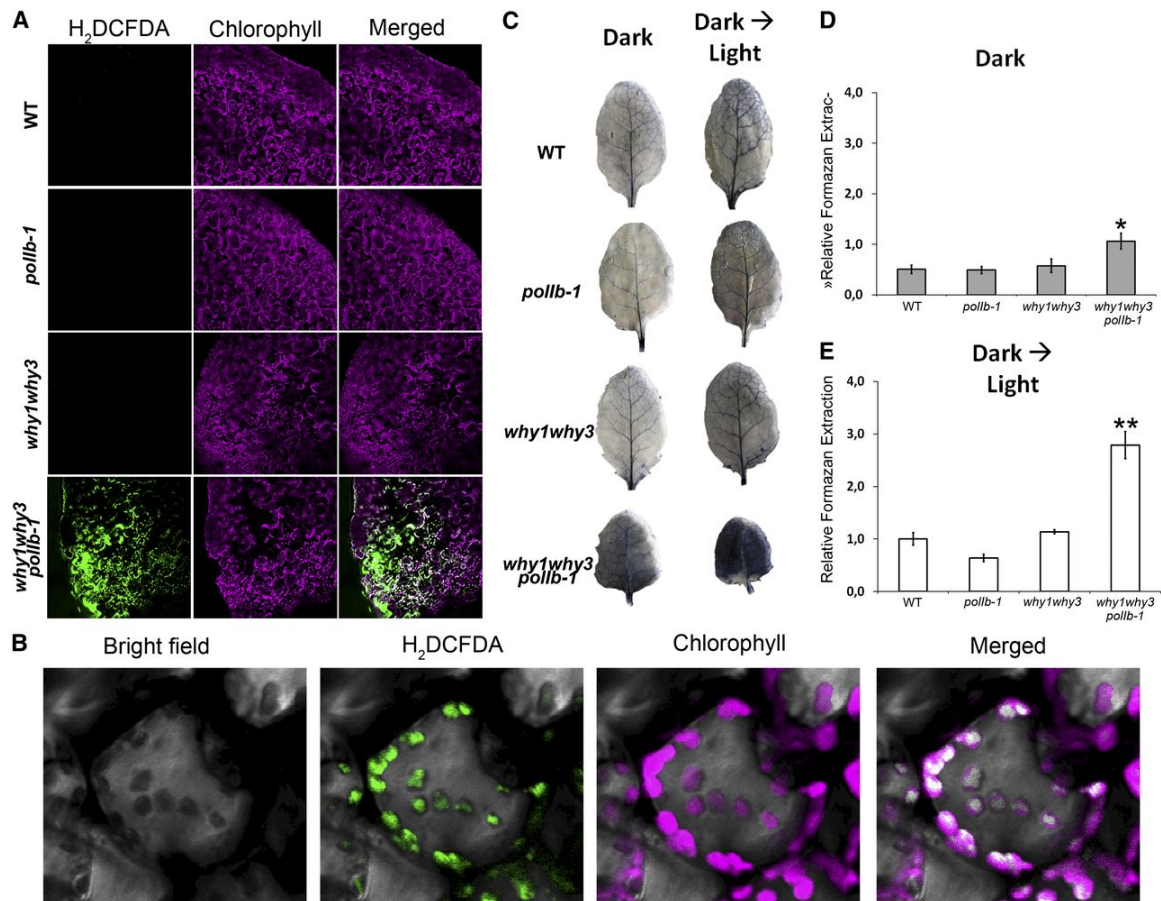


Figure 15. *why1why3pollb-1* Produces Elevated Levels of Chloroplast ROS in a Light-Dependent Manner. (A) Microscopic analysis of 28-day-old *Arabidopsis* leaves incubated with H₂DCFDA. Oxidized H₂DCFDA is represented in green, chlorophyll in magenta and the merge yields a white color. (B) Inset on H₂DCFDA foci, which colocalize with chloroplasts. (C) Representative photographs of NBT stained leaves from 28-day-old *Arabidopsis* plants. Blue color on the leaves is caused by formazan, the reduced form of NBT. (D) Histogram presenting the relative formazan concentration of the dark-incubated leaves. Five different leaves were tested for every condition. The error bars represent the standard deviation of the five extractions. One asterisk indicates a significant difference of a Student's t-test P-value ≤ 0.05 with the wild type. (E) Histogram presenting the relative formazan concentration of dark-incubated leaves switched to high light

for 15 minutes. Five different leaves were tested for every condition. The error bars represent the standard deviation of the five extractions. Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type.

To further characterize ROS production in *why1why3pollb-1*, detached leaves from four-week-old plants were incubated in the dark in the presence of the colorimetric agent nitroblue tetrazolium (NBT), which forms an insoluble quantifiable blue precipitate, the formazan, upon reduction from superoxide (O_2^-) [248]. While the color intensity of *pollb-1* and *why1why3* leaves was indistinguishable from that of wild type, *why1why3pollb-1* leaves showed a dense shadowed blue coloration all over their surface (Fig. 15C, Dark). Measurement of the formazan produced in those leaves revealed an approximate two fold increase in superoxide concentration as compared to WT leaves (Fig. 15D). ROS concentration was further analysed by switching the dark-incubated leaves to high light intensity. As expected, this treatment induced ROS production in the leaves of every line tested (Fig. 15C, Dark \rightarrow Light and 15E). However, the increase was more pronounced in the *why1why3pollb-1* mutant, further suggesting that the PET is involved in the redox imbalance of this mutant.

To verify that the elevated ROS levels in *why1why3pollb-1* are truly the consequence of the ptDNA rearrangements, ROS production was assessed in wild type plants treated either with CIP or NOVO. As depicted in Figure 16A, both treatments effectively affected plant development. Indeed, whereas the NOVO treatment arrested plant growth, the CIP treatment bleached the sides of the new emerging leaves (Fig. 16A). Treated plants were then incubated in presence of H_2DCFDA to assess the redox imbalance. Similarly to the *why1why3pollb-1* mutant, the CIP treatment also led to the formation of ROS foci that colocalized with chloroplasts at the subcellular level (Fig. 16B and 16C). No fluorescence could be observed in the NOVO-treated plants, suggesting that the redox imbalance observed is caused by ptDNA rearrangements, and not by gyrase inhibition.

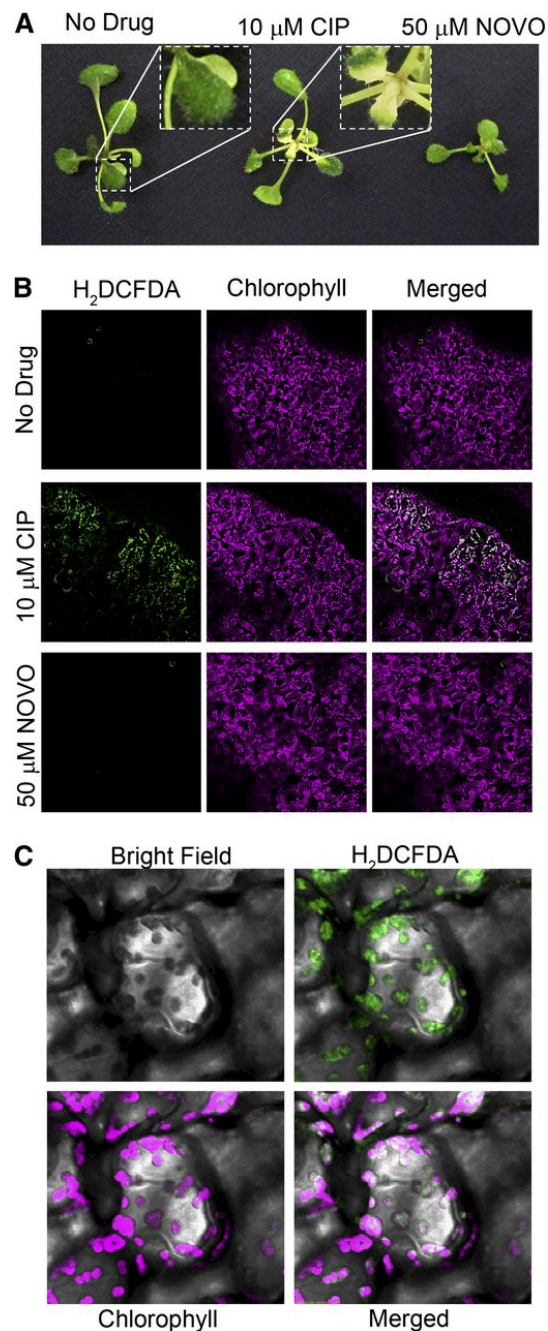


Figure 16. Ciprofloxacin-Treated Seedlings Show Signs of Bleaching Associated With Increased Chloroplast ROS Production. (A) Representative photographs of 15-day-old plants exposed from day 10 to 15 with the indicated concentration of drug. **(B)** Microscopic analysis of plants exposed to the prior treatment with H₂DCFDA. Photographs were taken on day 11, 24 hours after the beginning of the treatment. Oxidized H₂DCFDA is represented in green, chlorophyll in magenta and the merge yields a white color. **(C)** Inset on H₂DCFDA foci, which colocalize with chloroplasts.

2.4.3 The *why1why3pollb-1* Phenotype Depends on Elevated ROS Levels

Considering that the redox imbalance observed in *why1why3pollb-1* could be responsible for the appearance of the yellow-variegated phenotype, it is likely that decreasing the electron flow across the PET would rescue the phenotype. Therefore, to minimize the pressure on the electron transport chain, plants were grown for three weeks at low light and their phenotypes compared to those grown at normal light. When grown for three weeks under normal light conditions, *why1why3pollb-1* plants developed a severe dwarf and yellow-variegated phenotype (Fig. 14B and Fig. 17A, NL) that is associated with low chlorophyll content (Fig. 17B). In comparison, although all seedlings grown at low light for three weeks exhibited severe growth retardation, no additional growth delay and yellow variegation could be observed for *why1why3pollb-1* (Fig. 17A, LL 21 days). The rescue of the yellow-variegated phenotype persisted even after sixty days of growth under the same conditions, although at this stage the plants had smaller petioles than wild type and *why1why3* plants (Fig. 17A, LL 60 days). Consistent with the rescue of the phenotype, low light grown *why1why3pollb-1* plants exhibited chlorophyll content comparable to that of control plants (Fig. 17B). This low light dependent rescue of the phenotype thus suggests that elevated ROS levels play a role in the appearance of the *why1why3pollb-1* yellow-variegated phenotype. In agreement with this, *why1why3* and *why1why3pollb-1* plants are hypersensitive to paraquat and, therefore, affected by an aggravation of the redox imbalance (Fig. 27, page 89).

In contrast to the observed rescue of the phenotype, low light conditions did not restore PSII efficiency in *why1why3pollb-1* plants (Fig. 17C). This is in agreement with our hypothesis that the poor PET efficiency is due to ptDNA rearrangements, the abundance of which is also little affected by the light regimen (Fig. 28, page 90). Thus, it seems likely that ptDNA rearrangements do not originate from ROS-associated DNA modification. Consistent with this idea, the ROS-inducing agents paraquat and atrazine do not lead to an increased accumulation of MMBIR-associated ptDNA rearrangements (Fig. 29, page 91).

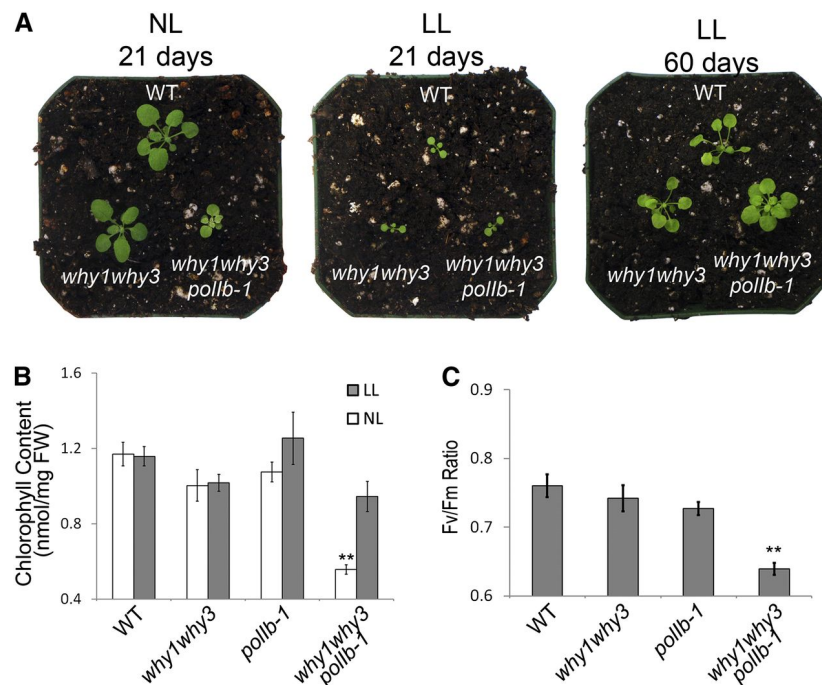


Figure 17. Growth under Low Light Conditions Rescues the *why1why3pollb-1* Phenotype. (A) Representative photographs of wild type, *why1why3* and *why1why3pollb-1* *Arabidopsis* plants grown under normal (NL) or low light (LL) regimens. (B) Chlorophyll content of each line grown under normal or low-light conditions for 21 and 60 days, respectively. White and Gray bars represent respectively NL and LL experiments. The error bars represent the standard deviation for five measurements on independent biological samples. Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type. (C) Fv/Fm ratios for 60-day-old plants of the indicated genotype grown under low-light conditions. The error bars represent the standard error of the mean (SEM). Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type. Every measurement was performed independently on at least five biological samples.

2.4.4 Yellow-Variegated Sectors in *why1why3pollb-1* Appear Photosynthetically Active and Are not Related to Cell Death

ROS readily damage all components of the cell and induce cell death if improperly scavenged [232-234]. To verify if the high ROS concentration in the *why1why3pollb-1* mutant is associated with cell death, four-week-old leaves were stained with trypan blue to selectively color dead cells [249]. Interestingly, although the *why1why3pollb-1* mutant produces high levels of ROS, cell death did not appear higher in leaves of this mutant than in leaves of wild

type plants (Fig. 30, page 92). This result suggests that the *why1why3pollb-1* yellow-variegated phenotype is not the direct consequence of ROS-associated cell death. To verify if chloroplasts in *why1why3pollb-1* appear functional, chloroplast ultrastructure was examined in green and yellow sectors of this mutant using transmission electron microscopy. Remarkably, despite the redox imbalance, thylakoid stacks appear normal in both green and yellow sectors of *why1why3pollb-1* (Fig. 18A). In comparison, chloroplast ultrastructure is completely altered in white variegated sectors of *why1why3* (Fig. 18A) [27]. Furthermore, whereas the white-variegated sectors of *why1why3* present no trace of chlorophyll, the yellow-variegated sectors of *why1why3pollb-1* are not devoid of photosynthetically active chloroplasts, as monitored by chlorophyll autofluorescence (Fig. 18B). Altogether, these results suggest that chloroplasts in *why1why3pollb-1* are still functional despite the redox imbalance.

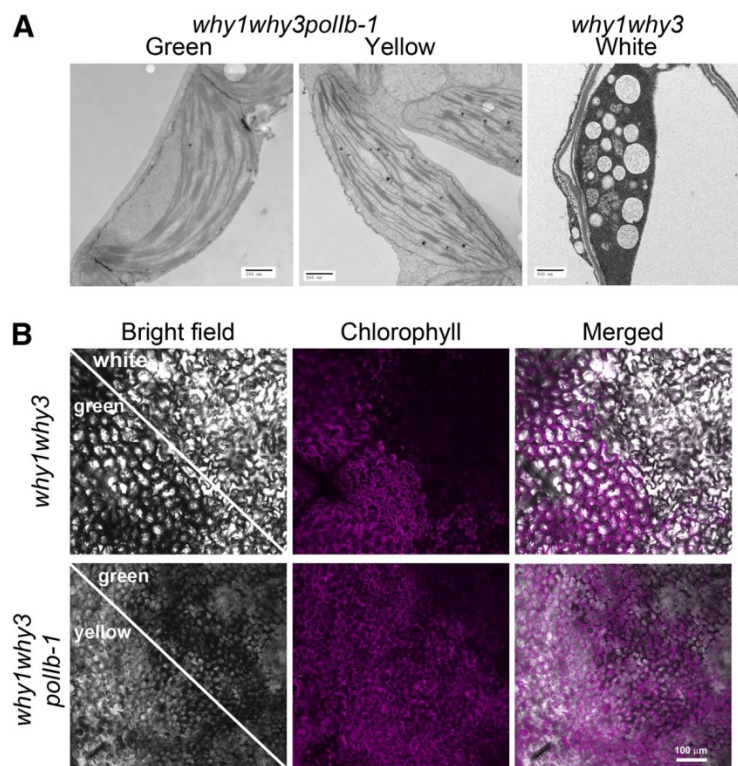


Figure 18. The *why1why3pollb-1* Yellow-Variegated Sectors Are not Devoid of Photosynthetically Active Chloroplasts. (A) Chloroplast ultrastructure observed by transmission electronic microscopy of green, white and yellow sectors of *why1why3* and *why1why3pollb-1* leaves. **(B)** Microscopic analysis of chlorophyll

autofluorescence in green, white and yellow sectors of 28-day-old *why1why3* and *why1why3pollb-1* leaves. Chlorophyll fluorescence is represented in magenta.

2.4.5 The *why1why3pollb-1* Yellow-Variegated Phenotype and Rearrangements Are not Maternally Inherited

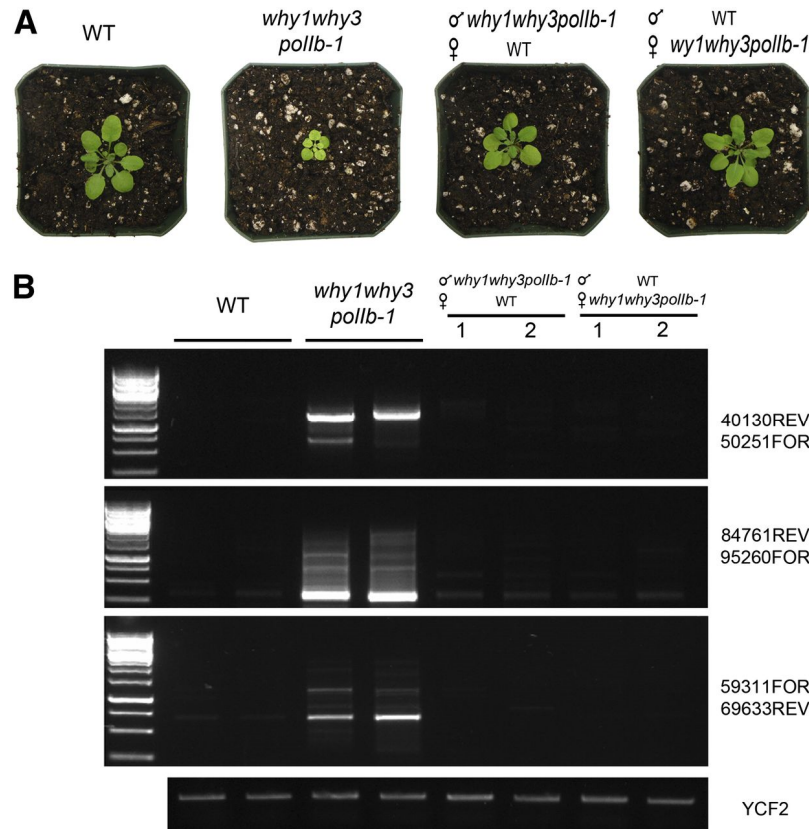


Figure 19. The *why1why3pollb-1* ptDNA Rearrangements and Phenotype Are not Maternally Inherited. (A) Representative photographs of 21-day-old backcrossed *why1why3pollb-1* plants. (B) Representative outward-facing PCR reactions showing the abundance of MMBIR-associated plastid DNA rearrangements of 14-day-old backcrossed *why1why3pollb-1* plants. PCR reactions were carried out using the indicated primer pairs. The amount of DNA loaded in each reaction was equilibrated relative to the amplification of a *YCF2* fragment. Every reaction was performed on three different biological samples and a representative biological duplicate is presented.

The appearance of variegation in many mutants has been linked to maternal transmission of “permanently-defective” chloroplasts [250]. However, the fact that chloroplasts in

why1why3pollb-1 appear functional and that the phenotype does not worsen in a generation-dependent manner suggests that its yellow-variegated phenotype is not maternally inherited. This was tested by backcrossing the *why1why3pollb-1* mutant with wild type plants and examining the transmitted phenotype. For each of the reciprocal crosses, 72 plants from five distinct crosses were examined. As expected, when *why1why3pollb-1* plants were backcrossed as male parents, no transmission of the phenotype was observed (Fig. 19A). Consistent with the idea that chloroplasts are functional in this mutant, the phenotype also disappeared in all plants issued from the reciprocal maternal crosses. Figure 19B indicates that ptDNA rearrangements were also eliminated in two-week-old plants of the F1 progeny, in agreement with the lack of inheritance of the phenotype in each reciprocal cross. Restoration of the alleles thus probably prevents generation of *de novo* ptDNA rearrangements in addition to allowing dilution of the inherited rearrangements throughout the entire plant. These results further support the idea that *why1why3pollb-1* chloroplasts function normally despite the redox imbalance.

2.4.6 *why1why3pollb-1* Exhibits a Photosynthetic Stress-Related Nuclear Genetic Reprogramming

We then tested whether the elevated ROS levels could lead to nuclear genetic reprogramming in *why1why3pollb-1*, enabling the plant to adapt to its particular photosynthetic conditions. To test this hypothesis, a RNA-Sequencing (RNA-Seq) approach was used to sequence and quantify the entire *why1why3pollb-1* nuclear transcriptome. Comparison of this transcriptome to that of wild type plants revealed a complex genetic reprogramming with 340 and 157 transcripts up and down-regulated at least two-fold, respectively (<http://www.plantphysiology.org/content/163/2/867/suppl/DC1>, Supplemental table S1).

To characterize gene expression reprogramming in *why1why3pollb-1*, its entire nuclear transcriptome was subjected to MapMan analysis and the biological processes with the most significant alterations were identified [251]. Regulation of transcription and RNA processing were the most significant terms, closely followed by photosynthesis and, particularly, the

light reactions (Fig. 20A). This was followed by metabolism, especially that of carbohydrate metabolism, which is not surprising considering that photosynthesis is affected in *why1why3pollb-1* (Fig. 14D). Among the other terms significantly overrepresented were the ascorbate and glutathione-associated redox processes as well as the abiotic stress responses. Interestingly, ethylene signal transduction was significantly overrepresented, similarly to a number of abiotic stresses that are also associated with redox imbalance [235,252,253].

To independently characterize genes up and down-regulated at least two fold in *why1why3pollb-1*, these two subsets of genes were analysed using Pageman [254]. Similarly to the MapMan analysis of the entire transcriptome, assignment of each of the up-regulated genes to a given biological process mainly resulted in the terms RNA, protein and signaling (Fig. 20B). Regarding the down-regulated genes, the terms hormone and stress were strongly represented in addition to genes of various functions, the miscellaneous class, which accounts alone for more than half of the genes. However, the main terms found using Pageman belongs to large classes containing large numbers of genes. To avoid this bias toward the largest classes, enrichment of each biological process was calculated relative to its class size (Fig. 20C). Interestingly, this normalization revealed a net enrichment for the terms hormone, stress, redox, nucleotide metabolism, RNA and signaling in the up-regulated genes. The down-regulated genes showed a clear enrichment for the terms photosynthesis, hormone and stress. Globally, a net enrichment for biological processes linked to photosynthetic stress is found within the most differentially expressed genes in *why1why3pollb-1*.

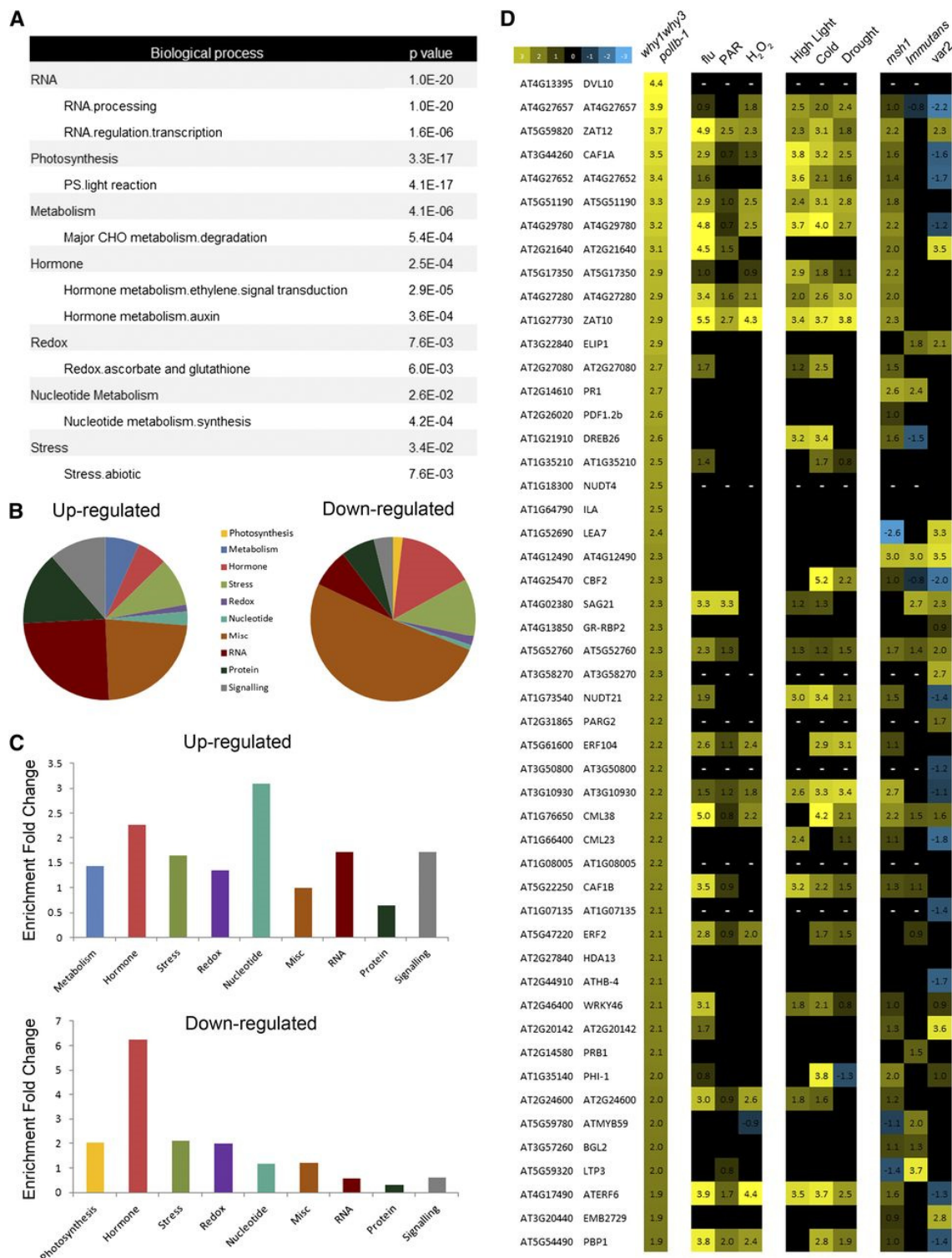


Figure 20. The *why1why3pollb-1* Mutant Exhibits a Complex Nuclear Genetic Reprogramming. (A) MapMan analysis for GO terms enrichment of the entire *why1why3pollb-1* nuclear transcriptome. (B) Pie charts presenting the biological processes associated to the significantly *why1why3pollb-1* up- and down-

regulated genes (2-fold). To lighten the Figure, the GO terms “unknown” and “not assigned” were removed from the analysis. **(C)** Histogram presenting the enrichment of each biological process of the *why1why3pollb-1* transcriptome. **(D)** Heatmap of the 50 most up-regulated *why1why3pollb-1* genes compared to other ROS-inducing conditions, abiotic stress responses and variegated mutants transcriptomes.

2.4.7 The *why1why3pollb-1* Nuclear Genetic Reprogramming Closely Resembles ROS, Abiotic Stress and *msh1* Responses

Numerous stresses such as chloroplast redox imbalance, abiotic stress and chloroplast biogenesis defects have been shown to initiate plastid-to-nucleus retrograde signaling pathways and modify the nuclear genetic expression patterns. To assess if the genetic reprogramming observed in *why1why3pollb-1* is similar to that observed in other conditions known to alter the nuclear transcriptome, the expression level of the fifty most overexpressed genes in this mutant was compared to that observed in other transcriptomic studies (Fig. 20D). The *why1why3pollb-1* transcriptome was first compared to that of other conditions known to produce specific types of ROS, including those of the *flu* mutant, which exhibits elevated levels of singlet oxygen upon a dark to light shift [96], of the treatment of Col-0 with paraquat, which produces superoxide in the PET [255,256], and of Col-0 treated with hydrogen peroxide. Interestingly, of the 43 overexpressed *why1why3pollb-1* genes found on the ATH1 chip (probes for seven of the *why1why3pollb-1* overexpressed genes are not present on this chip), 26, 18 and 15 were also overexpressed in *flu*, paraquat treatment and hydrogen peroxide treatment, respectively.

Several of the 50 most overexpressed genes in *why1why3pollb-1* encode key components of the abiotic stress responses, including the Early Light-Inducible Protein ELIP1, the transcriptional regulators ZAT10 and ZAT12, the ERF/AP2 transcription factor CBF2, and the mRNA deadenylases CAF1A and CAF1B [111,112,116,257-259]. Not surprisingly, a striking similarity was observed when comparison was made with overexpressed genes in abiotic stress responses. Indeed, 20, 26 and 22 genes were also up-regulated upon high light, cold

and drought stresses, respectively (Fig. 20D) [260,261]. A total of 15 genes that are overexpressed during all three abiotic stresses were also up-regulated in *why1why3pollb-1*.

Finally, since most variegated mutants are also linked to an alteration of the PET, we were interested in verifying the similitudes between the *why1why3pollb-1* transcriptome and those of the variegated mutants *msh1*, *var2* and *immutans* [220,262,263]. Although the *var2* and *immutans* mutants also destabilize the PET, only limited similarity was observed between their nuclear transcriptome and that of *why1why3pollb-1* (Fig. 20D). Only 12 of the 43 genes were up-regulated in *immutans*. Surprisingly, although 28 genes were differentially expressed in *var2*, 13 of them were repressed and thus represent a distinct genetic response than the one observed in *why1why3pollb-1*. Nevertheless, a strong correlation was observed with *msh1* since 31 genes were also up-regulated in this mutant. Our own analysis of the published *msh1* transcriptome revealed a high similitude with the *flu* mutant and the high light treatment. Indeed, 40% and 32% of the *msh1* two fold up-regulated genes are also overexpressed in *flu* and high light treatment, respectively. The fact that the MSH1 protein is involved in the maintenance of plastid and mitochondria genome stability [220] further supports the relationship between the observed transcriptome modifications and plastid genome rearrangements.

2.4.8 The *why1why3pollb-1* Nuclear Genetic Reprogramming Is not Observed Under Low-Light Conditions

The important signaling functions carried out by ROS in triggering stress responses have been extensively documented [96,100,237]. To assess if the nuclear genetic reprogramming observed in *why1why3pollb-1* depends on elevated chloroplast ROS production, we compared, in plants grown either under normal or low-light conditions, the expression of six genes whose expression is up-regulated in *why1why3pollb-1*. These genes are the Nudix hydrolase homolog 4 (*NUDT4*) [264], the oxidative stress marker *UPOX* (AT2G21640) [106,265], the ACC synthase 6 (*ACS6*) [266], the ARM repeat AT3G06530 and the two Early Light-Inducible Proteins (*ELIP1* and *ELIP2*) [116,121,267]. The overexpression of these *why1why3pollb-1* up-regulated genes was first confirmed by quantitative RT-PCR to validate

the RNA-Sequencing results. qRT-PCR confirmed that all these genes are significantly overexpressed in *why1why3pollb-1* (Fig. 21, NL). This overexpression was, however, not observed in *pollb-1* or *why1why3* (Fig. 31, page 93). Expression of these markers was then measured in plants grown for 21 days under low light conditions, where no difference in phenotype is observed between *why1why3pollb-1*, WT or *why1why3* (Fig. 17) and where ROS levels are low. Figure 21 (LL) shows that, under these conditions, there is no significant difference in the expression of these genes. These results again suggest that it is the ROS produced under normal light conditions that act as signals to initiate the genetic reprogramming in response to the *why1wh3pollb-1* ptDNA rearrangements.

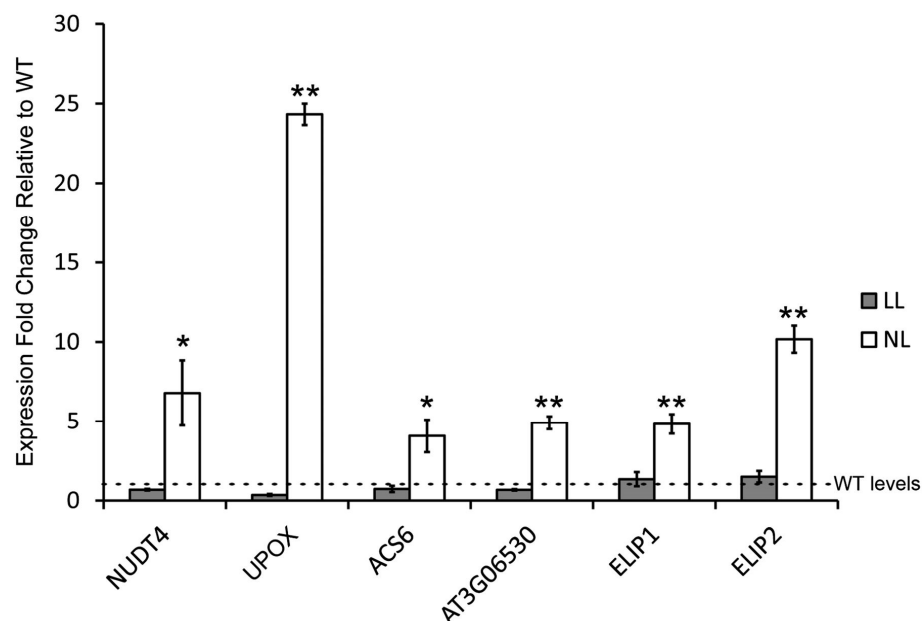


Figure 21. The *why1why3pollb-1* Nuclear Genetic Reprogramming Is Not Observed under Low-Light Conditions. Histograms presenting the real-time PCR measurements of *NUDT4*, *UPOX* (AT2G21640), *ACS6*, *AT3G06530*, *ELIP1* and *ELIP2* expression in *why1why3pollb-1* relative to wild type plants for normal light (NL) and low light (LL) conditions. The error bars represent the standard error of the mean (SEM) of the biological triplicates. One and two asterisks indicate, respectively, a significant difference of a Student's t-test P-value ≤ 0.05 and 0.01 with the wild type.

2.4.9 Genetic Reprogramming in *why1why3pollb-1* Allows Adaptation to High Light Stress

Reprogramming of the *why1why3pollb-1* nuclear transcriptome suggests that this mutant line could adapt to elevated ROS concentrations. This was verified by growing plants for three weeks at normal light, followed by transfer to high light conditions for an additional week. Remarkably, the *why1why3pollb-1* plants were considerably less affected by this treatment than the wild type, *why1why3* or *pollb-1* plants (Fig. 22A). Under these conditions, the *why1why3pollb-1* plants accumulated about 20 times less anthocyanin in their leaves than control plants (Fig. 22C). The greater tolerance to high light in addition to the reduced accumulation of the antioxidant anthocyanin in response to this treatment thus supports an acquired adaptation to elevated ROS levels in *why1why3pollb-1*.

To confirm that this adaptation to high light conditions is acquired rather than innate, plants were germinated and allowed to grow for four weeks under constant high light conditions. The lack of pre-adaptation under normal light conditions abolished the tolerance of *why1why3pollb-1* plants to high light, as these mutants were considerably more affected than control plants (Fig. 22B). Expression pattern of the high light and oxidative stress markers *ELIP1* and *ELIP2* supported that an acquired adaptation is taking place in this mutant [116,121,267]. Indeed, when measured in 5-day-old seedlings grown under normal light, expression levels of *ELIP1* and *ELIP2* were low and did not differ in wild type and *why1why3pollb-1* plants (Fig. 22D). Their expression was however significantly higher than in wild type plants after 14 days in normal light. As these genes are known to be induced by ROS, these results further support the idea that modification of the nuclear expression and adaptation to subsequent oxidative stresses are processes that first require an early accumulation of ROS.

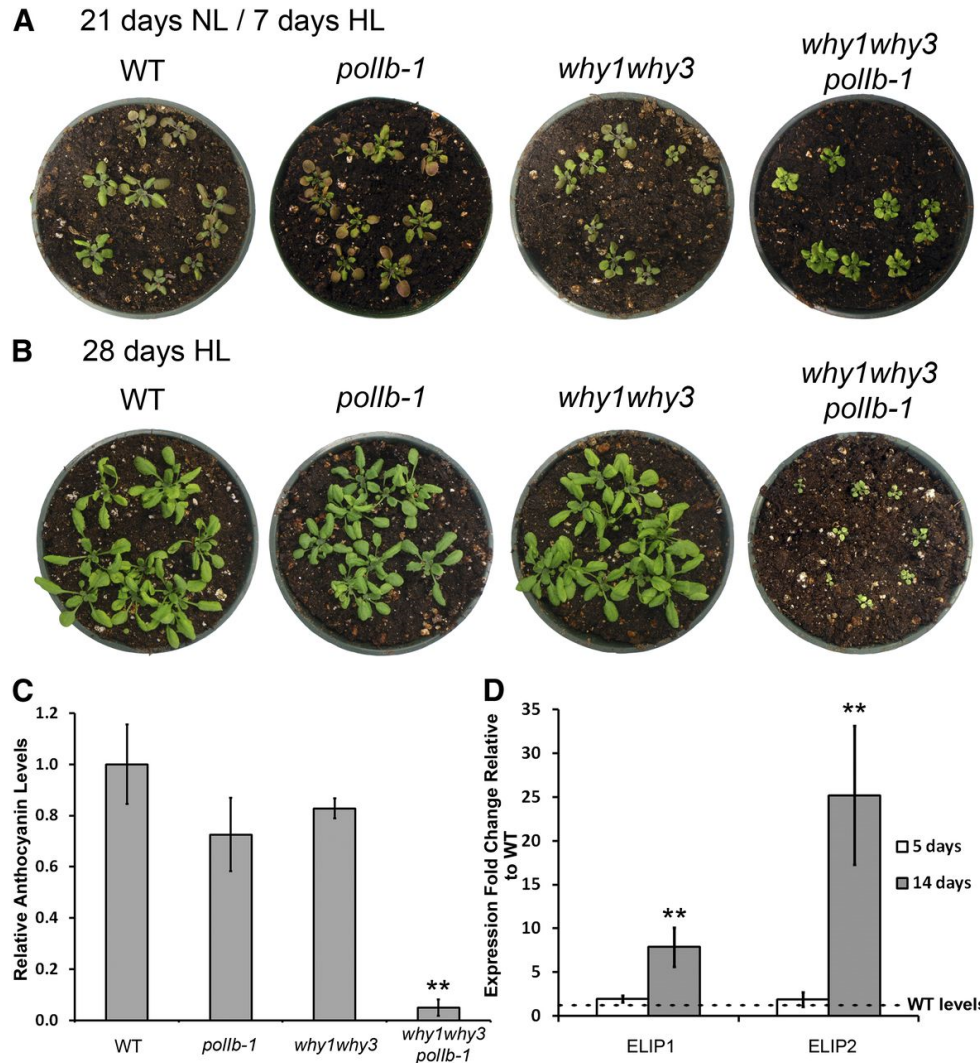


Figure 22. The Early Oxidative Stress in *why1why3pollb-1* Leads to Acquired Resistance to High Light. (A) Photographs of representative *Arabidopsis* plants grown for 21 days at normal light (NL) and switched to high light (HL) for an additional seven days. (B) Photographs of representative *Arabidopsis* plants grown for 28 days at high light (HL). (C) Histogram presenting the relative anthocyanin content for the mutant lines and wild type plants switched for seven days at high light. The error bars represent the standard deviation of the biological triplicates. Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type. (D) Histogram presenting the real-time PCR measurements of the *ELIP1* and *ELIP2* expression in *why1why3pollb-1* relative to wild type plants grown under normal light conditions for 5 and 14 days. The error bars represent the standard error of the mean (SEM) of the biological triplicates. Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with the wild type.

2.5 Discussion

2.5.1 Plastid Genome Instability Leads to Impairment of Photosynthesis and Increased Chloroplast ROS Production

Until now, few mutations in the plastid genome have been associated to photosynthetic impairment and lethality [268], and the impact of more generalized plastid genome instability is still largely unknown. In this article, we used the *why1why3pollb-1* mutant line and the organelle specific DSB-inducing agent, CIP, to analyse the consequences of generalized plastid genome instability.

First, the *why1why3pollb-1* mutations and CIP treatment were both shown to reduce plant photosynthetic capacity, as indicated by the reduction of the Fv/Fm ratio (Fig. 14D and 26, page 88). The effect of CIP is most likely due to its DSB-inducing activity as NOVO, another gyrase inhibitor that does not induce DSBs, does not influence the Fv/Fm parameter. One interpretation of these results is that, since many ptDNA genes encode components of the PET, ptDNA rearrangements will invariably affect these genes, leading to a modification of the stoichiometry of the components of the PET. The fact that an increased chloroplast ROS production could only be observed in the *why1why3pollb-1* mutant and the CIP-treated plants supports this link between ptDNA rearrangements and severe photosynthetic defects (Fig. 15 and 16). Moreover, as shown by the NBT staining experiment, the abundance of superoxide in *why1why3pollb-1* is strongly and quickly enhanced by light, which further supports an imbalance at the level of the PET [269]. Interestingly, in mammals, a similar redox imbalance was also observed in mitochondria, where genome rearrangements are linked to oxidative phosphorylation defects and ROS production [270,271].

The plastid genome has long been known to be more stable than its nuclear counterpart [20], but the reason for such a high stability remains unknown. The developmental defects and redox imbalance observed in *why1why3pollb-1* and upon CIP treatment demonstrate the deleterious consequences associated with plastid genome instability. Indeed, CIP-treated seedlings exhibit bleaching of their new emerging leaves and *why1why3pollb-1* plants

present a yellow-variegated dwarf phenotype. Consistent with this finding, the severity of the *why1why3pollb-1* phenotype correlates with the intensity of chloroplast ROS production (Fig. 17, 22B and 27, page 89). All these deleterious effects associated with plastid genome instability might explain the importance of maintaining its highly optimized genome structure.

2.5.2 ptDNA Rearrangements in *why1why3pollb-1* Do not Lead to “Permanently-Defective” Chloroplasts

Variegation has been studied for half a century, and the mode of inheritance of almost every variegated mutant identified to date has been established. While some variegated mutants show maternal inheritance, others are inherited in a Mendelian fashion [250]. It has been proposed that in the case of maternally inherited variegation, a perturbation would lead to “permanently-defective” chloroplasts. In *Arabidopsis* and most angiosperms, these “permanently-defective” chloroplasts are then maternally transmitted to the progeny, with the extent of variegation usually correlating with the proportion of transmission. In *why1why3pollb-1*, chloroplasts appear to be photosynthetically active and their ultrastructure is maintained. Accordingly, the yellow variegation observed in this mutant is inherited in the classical Mendelian way, supporting the idea that the chloroplasts are not “permanently-defective”. Conversely, since white variegation appears in *why1why3pollb-1* at the same frequency as in the *why1why3* population, it suggests that both phenotypes are actually the same. Thus, the white variegated sectors of *why1why3pollb-1* most likely possess “permanently-defective” chloroplasts with an altered ultrastructure.

The early studies on ptDNA inheritance were done with lines carrying point mutations at or near homoplasmy in their plastid genome [272,273]. In such cases, the transmission of the mutation is directly linked to the transmission of the chloroplast. In *why1why3pollb-1*, the progeny from the reciprocal crosses with wild type plants shows wild type levels of DNA rearrangements. This suggests that the rearranged molecules are present in heteroplasmy and relatively low copy numbers, which results mainly in the transmission of normal copies of the plastid genome. Once again the fact that both the phenotype and the rearrangements

are not maternally inherited supports the importance of these DNA rearrangements for the appearance of the yellow-variegated phenotype. In addition, restoration of the alleles in the crossing experiments prevents the formation of *de novo* ptDNA rearrangements and favors the dilution of potentially inherited rearrangements. Furthermore, the *why1why3pollb-1* phenotype remains constant through many generations, which suggests that most DNA rearrangements observed in its plastid genome are generated during plant development rather than inherited from the female gametes.

In *why1why3* plants, rearrangements throughout the plastid genome are also observed, although at a much lower frequency than in *why1why3pollb-1*. In approximately 5% of the *why1why3* plants, the amplification of specific parts of the genome, probably near homoplasmy, leads to the appearance of white variegated sectors [27]. The progeny of these variegated plants is composed of a large fraction of variegated plants (between 20% and 50%) [27]. Interestingly, the white variegated sectors in these plants show abnormal cell morphology and absence of chlorophyll autofluorescence. One could thus assume that the specific amplification causing the white-variegation leads to “permanently-defective” chloroplasts. White variegation is also observed in an equivalent fraction of the *why1why3pollb-1* population, which suggests that these two phenotypes are distinct and independent.

2.5.3 The ROS-Associated *why1why3pollb-1* Nuclear Genetic Reprogramming Promotes Acquired Resistance to High Light

ROS can have various deleterious effects on plant cell development. Therefore, plants have evolved mechanisms to counteract this threat. However, ROS are also important signaling molecules that ensure the response to biotic and abiotic stresses [274,275]. We propose that in *why1why3pollb-1*, the elevated chloroplast ROS levels are perceived by the cells and lead to genetic reprogramming, as indicated by the transcriptome modifications observed in this mutant. The high similitude of this genetic expression remodelling with that of many ROS-inducing conditions (Fig. 20D) suggests that ROS, and not ptDNA rearrangements, mediate this response. This is further supported by the apparent absence of genetic reprogramming

in low light grown *why1why3pollb-1* plants (Fig. 21) where ptDNA rearrangements are still present (Fig. 28, page 90). At this time it is however not possible to completely rule out that the excitation pressure, which is also influenced by light, contributes to the observed genetic reprogramming [276].

Surprisingly, no genes involved in the maintenance of plastid genome integrity were differentially expressed in *why1why3pollb-1*. This is rather unexpected as preventing ptDNA rearrangements could prevent ROS formation and its damages to cellular components. Chloroplast ROS therefore do not appear to act as direct signals of genome perturbation but rather provide a warning that the PET has been affected and that the plant must adapt to stress. This could explain the striking resemblance observed between the genetic reprogramming in *why1why3pollb-1* and the responses to abiotic stresses (Fig. 20D). Examples of this are the early light induced proteins *ELIP1* and *ELIP2*, often considered as markers of high light stress [116,121,267], which are significantly overexpressed in *why1why3pollb-1*. Overexpression of the abiotic-stress-responsive transcriptional activators ZAT10, ZAT12, CBF1 and CBF2 also supports this hypothesis [111,112,257,258]. Finally, a large proportion of the overexpressed *why1why3pollb-1* genes were shown to be involved in calcium and ethylene signaling pathways. The important roles mediated by these pathways in the regulation and signalization of abiotic stresses suggest that these molecules might also be important for proper adaption to plastid genome instability [277-279]. This adaptation appears to take place between days 5 and 14 (Fig. 22D) and promotes acquired resistance to high light. Although the redox imbalance observed in *why1why3pollb-1* is associated with deleterious phenotypes, the early oxidative stress experienced in this mutant might trigger this acquired resistance. This type of adaptation has also been reported for other chloroplast stresses and is generally referred to as the acclimation response [104,280-283]. Indeed, many genes differentially expressed in *why1why3pollb-1* were also shown to be part of the high light systemic acquired acclimation (HL SAA) response [283] and support that the *whywhy3pollb-1* nuclear reprogramming would be at the origin of the acquired tolerance to high light observed in this mutant. Interestingly, a similar high light tolerance was reported in

msh1, but the exact cause of this adaptation remains elusive, as both chloroplast and mitochondria exhibit genomic instability in this mutant [220]. Similarly to Msh1, PolIB is known to act in both organelles [43,244] and thus it is possible that alterations in DNA metabolism from both chloroplast and mitochondria contribute to the nuclear genetic reprogramming observed in *why1why3polib-1*. However, the synergistic increase in ptDNA rearrangements observed when plastid-localized Whirly proteins and PolIB are affected suggests that chloroplast defects are the main factor at the origin of changes in the nuclear gene expression.

Taken together, our results indicate that DNA rearrangements in the chloroplast genome lead to a severe PET imbalance and an important increase in photosynthesis-dependent ROS production. This elevated ROS production would then trigger plastid-to-nucleus retrograde signaling pathways that, in turn, reprogram the nuclear transcriptome. Although the *why1why3polib-1* mutant exhibits a dwarf and yellow-variegated phenotype, this reprogramming would allow the plants to survive to its inner oxidative stress and even acquire a high light tolerance.

2.6 Materials and methods

2.6.1 Plant Material and Growth Conditions

Arabidopsis (*Arabidopsis thaliana*; ecotype Columbia-0) mutant lines *polib-1* (SALK_134274), *why1why3* and *why1why3polib-1* were reported previously [27,43]. Seeds were sown on soil, vernalized for 3 days at 4°C and grown either under laboratory normal light (100 $\mu\text{mol m}^{-2} \text{s}^{-1}$), very low light (20 $\mu\text{mol m}^{-2} \text{s}^{-1}$) or moderate high light conditions (225 $\mu\text{mol m}^{-2} \text{s}^{-1}$) at 22°C on a 16 h day/8 h dark cycle. When grown at very low light, plants were germinated at 10 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for one week prior to their transfer at 20 $\mu\text{mol m}^{-2} \text{s}^{-1}$. For paraquat and atrazine experiments, seeds were sterilized and sown on Murashige and Skoog basal media (MS medium, Sigma-Aldrich) supplemented with 1 % sucrose, 0.8 % agar and the indicated amount of drug. For ciprofloxacin and novobiocin treatments, seeds were grown on solid MS

medium for eight days and transferred in liquid MS medium for two days to allow adaptation to the new growth conditions. The indicated amount of drug was then added at day 10.

2.6.2 Detection of ptDNA Rearrangements

Total DNA was isolated for each plant sample using a cetyl trimethylammonium bromide DNA extraction protocol [249]. For every DNA samples, ptDNA amount was equilibrated with a low-cycle amplification of a *YCF2* DNA fragment. Primers and approach for rearrangements detection were described previously [29] and PCR products were visualized after migration on GelRed (Biotium) stained agarose gels. Every reaction shown is representative of the eight primer pairs presented in Table I (page 94).

2.6.3 Measurement of PSII Efficiency

For each line analysed, Fv/Fm measurements were performed on whole, four-week-old plants grown under normal light conditions or whole, 60-day-old plants grown under low light conditions with an Open FluorCam (Photon Systems Instruments) according to the manufacturer's instructions.

2.6.4 Chlorophyll Extraction and Quantification

For each line analyzed, 10 mg of three-week-old plants grown under normal light conditions or 60-day-old plants grown under low light conditions were grinded in 400 μ L dimethylformamide and centrifuged for one minute at 13000 RPM. The supernatant was recovered and absorbance was measured at A_{645} and A_{663} . Chlorophyll content was then assessed as described previously [284].

2.6.5 Chloroplast Ultrastructure Imaging

Imaging of *why1why3pollb-1* chloroplast ultrastructure was performed in dissected yellow and green sectors as described previously [27].

2.6.6 ROS Detection, Quantification and Microscopic Analyses

The NBT superoxide coloration experiment was conducted as described previously [248]. Briefly, petioles of 28-day-old *Arabidopsis* leaves were cut twice under water to avoid air obstruction of the vasculature. For probe feeding of the leaves, petioles were soaked in the dark for 2 hours in 6 mM nitrotetrazolium blue chloride (Sigma-Aldrich) dissolved in 10 mM potassium phosphate pH 7.8. To assess the effect of photosynthesis on superoxide production, fed leaves were transferred at high light for fifteen minutes. To remove chlorophyll, colored leaves were boiled for five minutes in an ethanol, lactic acid and glycerol solution (4:1:1, v/v/v). To quantify formazan, leaves were boiled in dimethyl sulfoxide until leaves were clear and formazan concentration was measured spectrophotometrically at A560. Absorbance was measured in five leaves for every line and conditions tested and then normalized relative to the weight of each leaf. For the H2DCFDA experiment, protocol was adapted from a previous report [285]. Briefly, 10 day-old seedlings, acclimated for 48 hours in liquid MS medium, were incubated in 10 μ M H2DCFDA (Sigma-Aldrich) in PBS buffer for 30 minutes under high light exposure and washed twice for 5 minutes in PBS buffer. Chlorophyll autofluorescence and oxidized H2DCFDA were visualized using an Olympus FV300 confocal microscope. Chlorophyll and H2DCFDA were excited with a 488 nm argon laser. H2DCFDA fluorescence was detected at 535 nm and chlorophyll at 650 nm and over.

2.6.7 Anthocyanin Extraction and Quantification

For each line analyzed, three-week-old plants grown under normal light conditions were transferred for one week at high light. Anthocyanin extraction was then conducted as described previously [286] for five 30 mg replicates of each line.

2.6.8 RNA Isolation, RNA-Sequencing and Quantitative RT-PCR Analysis

For each plant sample total RNA was isolated using TRIzol reagent (Invitrogen) and the integrity of every sample assessed using a Bioanalyzer platform (Agilent). Libraries were prepared using the TruSeq RNA Sample Preparation v2 kit (Illumina) according to the manufacturer's instructions. Samples were then pooled together and ran on an Illumina

MiSeq to further verify the libraries quality. Sequencing was performed using the Illumina HiSeq2000 machine running TruSeq SBS v3 chemistry at the genomic platform of the Institute for Research in Immunology and Cancer, Université de Montréal. Cluster density was targeted at around 600-800k clusters/mm². Analysis was then carried out using the Galaxy online software suite [287]. Transcripts with FPKM values lower than one were discarded from the analysis to ensure optimal quantification. MapMan was used to retrieve GO annotations associated with the up- and down-regulated genes of the entire *why1why3pol1b-1* transcriptome [251]. Pageman web browsing tool was used to retrieve the GO annotations associated with the significantly up- and down-regulated *why1why3pol1b-1* genes [254]. Genevestigator was used to compare the 50 most overexpressed *why1why3pol1b-1* genes to that of other conditions [288]. Genevestigator transcriptomes for paraquat, hydrogen peroxide and drought treatments were, respectively, AT-120:Kudla/Puchta/Bartels/Harter/Nover, AT-185:Mittler and AT-58:Warren.

For quantitative RT-PCR analysis, cDNA synthesis was carried out using a first-strand cDNA synthesis kit (Fermentas) on RNAs from 14 and 21 day-old seedlings grown under normal and low-light conditions, respectively. cDNAs were diluted 20-fold prior to qRT-PCR experiments. Primers used for qRT-PCR experiments were designed for randomly-selected intron-containing genes and the 6 primer pairs with acceptable efficiency were selected (efficiency between 1.90 and 2.05). These primers are listed in Table II (page 94). Every reaction was carried out on biological and technical triplicates and normalized relative to β -Tubulin. Real-time qPCR reactions were done using the SYBR green master mix (SABiosciences, www.sabiosciences.com) according to the manufacturer's instructions. Melting curves confirmed the amplification of a unique PCR product for each of the qRT-PCR primer pairs used. LightCycler480 (Roche) was used for qRT-PCR experiments and data analyzed using the Light-Cycler480 software version 1.5.

2.6.9 Accession Numbers

Sequence data from this article can be found under the following accession numbers: WHY1 (AT1G14410), WHY3 (AT2G02740), POL1B (AT3G20540).

2.7 Acknowledgements

The assistance of Jean-Benoit Charron, Marianne Arteau and Sébastien Truche with the Fv/Fm, RNA-Sequencing and microscopy experiments, respectively, is gratefully acknowledged. We gratefully acknowledge Samuel Tremblay-Belzile for reading the manuscript.

2.8 Author contributions

E.L., E.Z. and N.B. designed research. E.L. and E.Z. performed research. E.L., E.Z. and N.B. analyzed the data. E.L., E.Z. and N.B. wrote the article.

2.9 Supplementary data

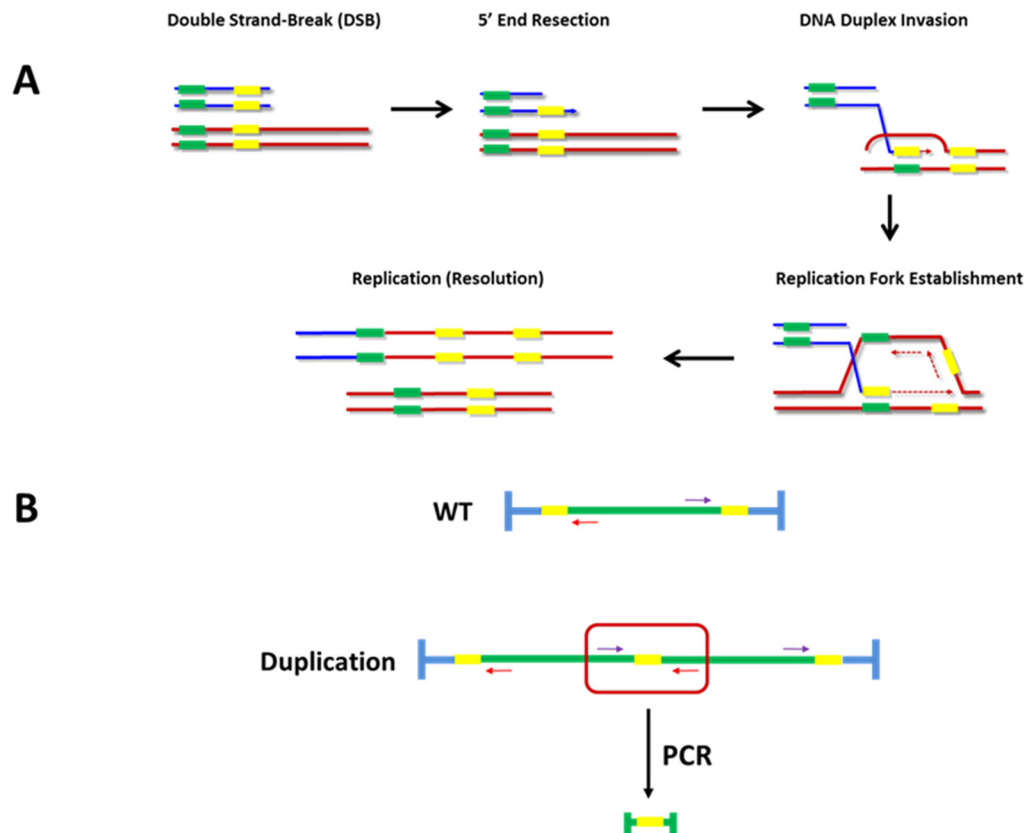


Figure 23. Microhomology-Mediated Break-Induced Replication (MMBIR) Leads to DNA Rearrangements. (A) Scheme presenting the repair of a DNA double-strand break by MMBIR. Dotted lines represent newly synthesized DNA. Yellow and green boxes represent microhomologous repeat sequences. **(B)** Scheme presenting the detection of MMBIR-associated DNA rearrangements by outward-facing PCR. Arrows represent the position of the PCR primers.

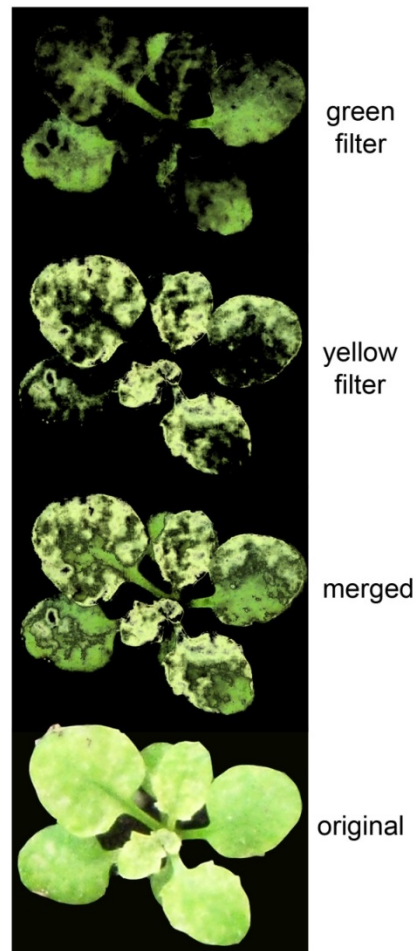


Figure 24. The *why1why3pollb-1* Mutant Develops Distinct Yellow and Green Sectors on Its Leaves. Photographs of a representative 28-day-old *why1why3pollb-1* yellow variegated plant. From top to bottom, the images represent the use of a green filter, a yellow filter, a merge of the two filters and the original picture. Color filters were applied using Photoshop CS4 color range function.



Figure 25. Some *why1why3pollb-1* Plants Exhibit Yellow and White Variegation Simultaneously. Photograph of a representative 28-day-old *why1why3pollb-1* yellow and white variegated plant.

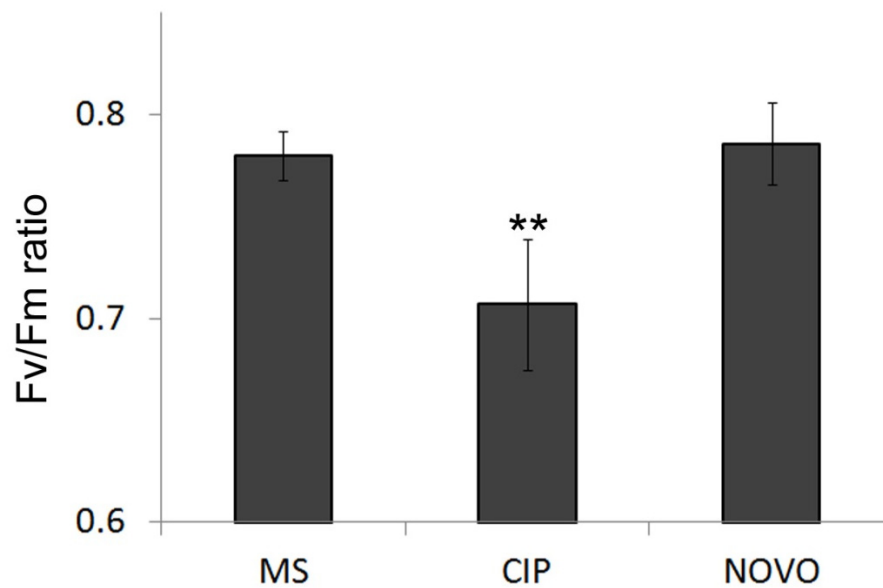


Figure 26. Ciprofloxacin Negatively Impacts Photosynthetic Efficiency in *Arabidopsis*. Fv/Fm ratios of 15-day-old plants treated from days 10 to 15 to 50 µM NOVO, 10 µM CIP or no drug (MS). The error bars represent the standard deviation. Two asterisks indicate a significant difference of a Student's t-test P-value ≤ 0.01 with MS. Measurements were performed independently on at least five biological samples.

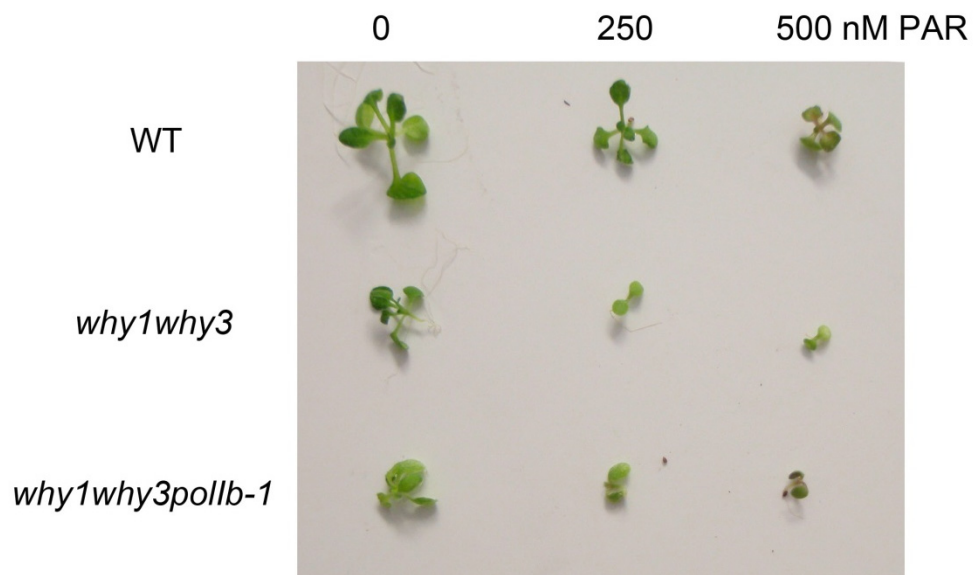


Figure 27. *why1why3* and *why1why3pollb-1* Plants Are Hypersensitive to Paraquat. Representative photograph of 3-week-old plants grown under normal light conditions on MS basal media containing 0, 250 or 500 nM paraquat (PAR).

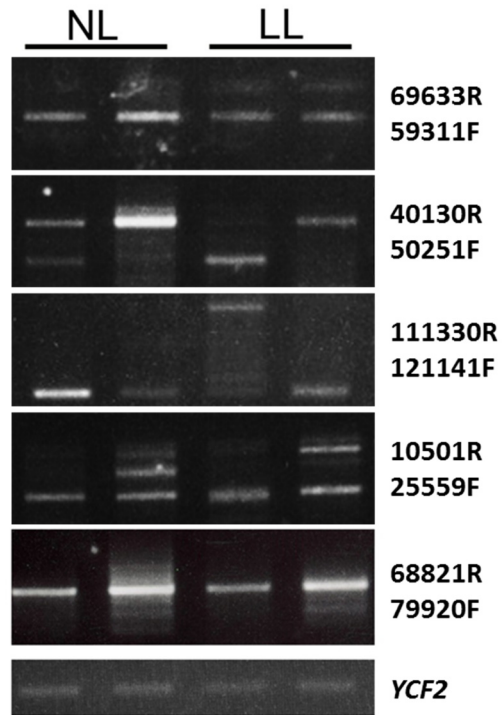


Figure 28. The Abundance of MMBIR-Associated ptDNA Rearrangements Is Little Affected by the Light-Regimen in *why1why3pollb-1*. Representative outward-facing PCR reactions showing the abundance of MMBIR-associated plastid DNA rearrangements of 3-week-old *why1why3pollb-1* plants grown either under normal (NL) or low-light conditions (LL). PCR reactions were carried out using the indicated primer pairs. The amount of DNA loaded in each reaction was equilibrated relative to the amplification of a *YCF2* fragment. Every reaction was performed on three different biological samples and a representative duplicate is presented.

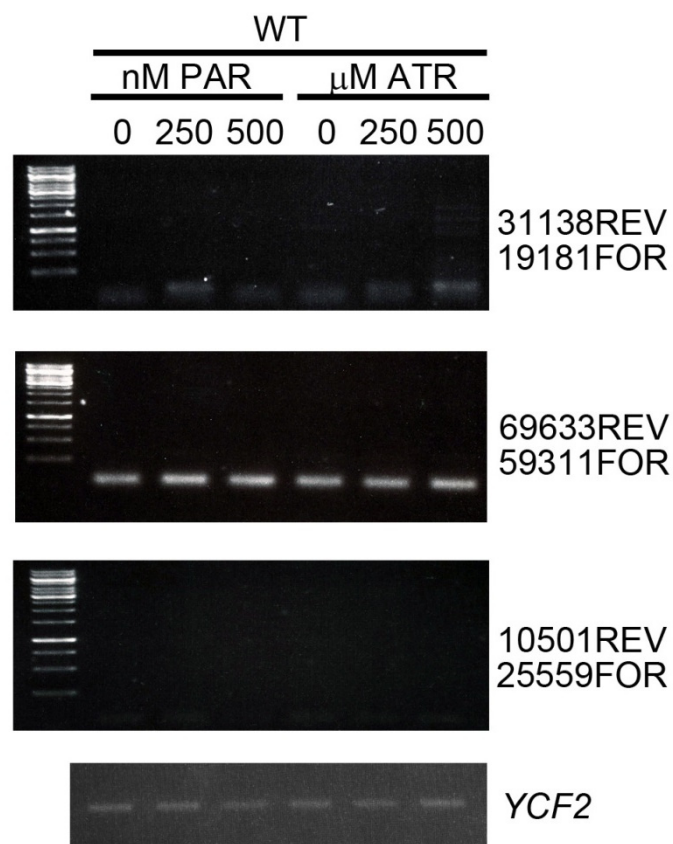
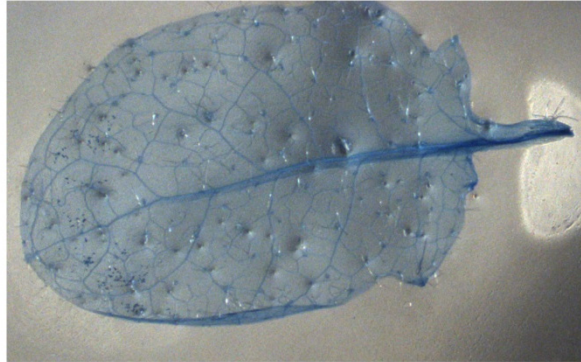


Figure 29. Paraquat and Atrazine do not induce MMBIR-Associated ptDNA Rearrangements. Representative outward-facing PCR reactions showing the abundance of MMBIR-associated plastid DNA rearrangements of 3-week-old *Arabidopsis* plants grown in the presence of the ROS-inducing agents paraquat (PAR) and atrazine (ATR). PCR reactions were carried out using the indicated primer pairs. The amount of DNA loaded in each reaction was equilibrated relative to the amplification of a *YCF2* fragment. Every reaction was performed on three different biological samples and a representative reaction is presented. The results indicate that MMBIR-associated ptDNA rearrangements are not the consequence of superoxide, singlet oxygen or hydrogen peroxide oxidation.

WT



why1why3pollb-1

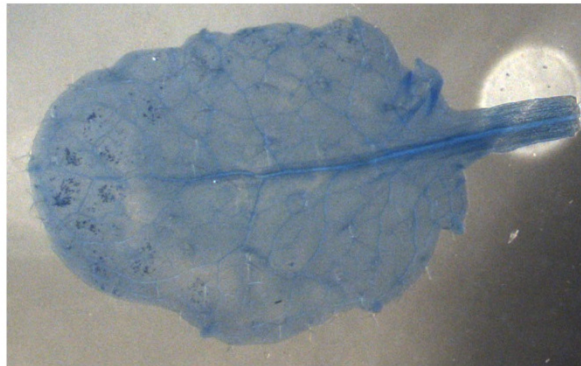


Figure 30. Elevated ROS Levels Do not Lead to Cell Death in *why1why3pollb-1*. Representative photographs of 4-week-old plants grown under normal light conditions and stained with trypan blue.

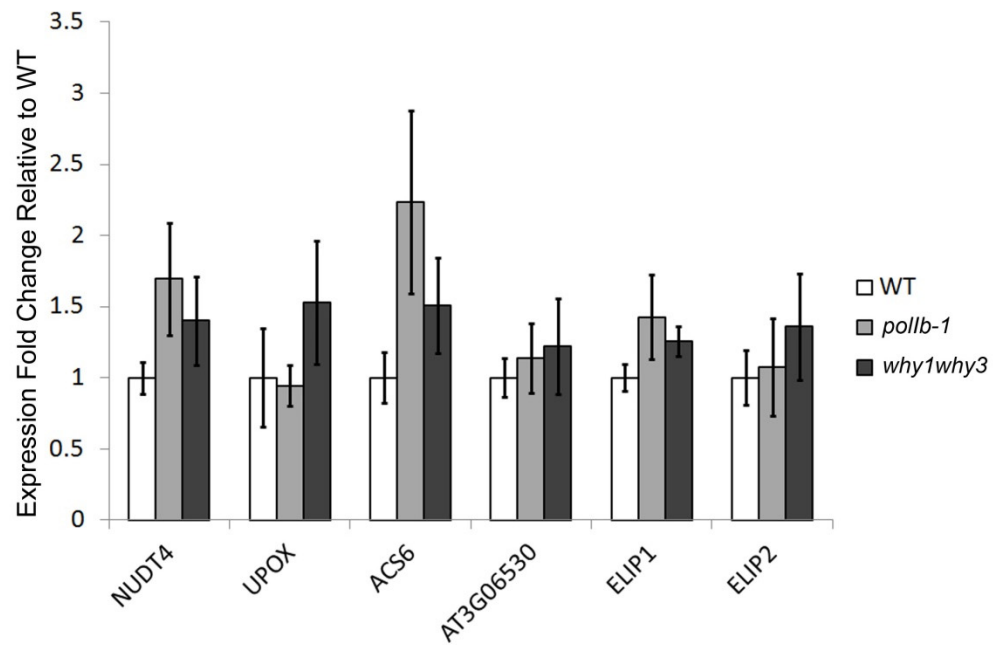


Figure 31. The *why1why3pollb-1* Nuclear Genetic Reprogramming Is Not Due to the *pollb-1* and *why1why3* Genetic Backgrounds. Histograms presenting the real-time PCR measurements of *NUDT4*, *UPOX* (AT2G21640), *ACS6*, *AT3G06530*, *ELIP1* and *ELIP2* expression in wild type, *pollb-1* and *why1why3* plants. The error bars represent the standard error of the mean (SEM) of the biological triplicates.

Tableau I. Primer Pairs Used for the Specific Amplification of MMBIR-Associated ptDNA Rearrangements. Primers are described previously in [29].

Primer number	Primer name	Sequence
1	31138REV	AAT AGC CTA GCG CAT CGA ACG
	19181FOR	CAG CAT CTG ATG TGC GTA CAG
2	40130REV	GGG AAG AGG AGG AAC ATG TCA
	50251FOR	GGC CCC TTC CTA ATT GGA GAT
3	84761REV	CCT TAG GCA CGG CCA TAC ATA
	95260FOR	AAC CTG CTA GTG GAG GAA GAC
4	105961REV	GCG TTA CTC AAG CCG ACA TTC
	116007FOR	AAA CCA CTC ATG CCC GGC AAT
5	69633REV	GAG GAA CAT GTA CTC GGG TGT
	59311FOR	AGG CCT AGT CTT TCC GGG AAT
7	50461REV	GAG ACG TCT ACC ACA ACA ACC
	59311FOR	AGG CCT AGT CTT TCC GGC AAT
8	68821REV	ACG CGG ATA TAT AGC TCG GAG
	79920FOR	CCA TGC CGT TCG GAA GTG ATT
9	10501REV	TGA CTG CCT TAC CAA TCG TCG
	25559FOR	CGG CGA CCA ATC CTT CCT AAT
ycf2	ycf2REV	TGT TTC GCC TCT TAC TCG GAG
	ycf2FOR	GAT CTC TGA GAG CTG TTT CCG

Tableau II. Primer Pairs Used for Quantitative RT-PCR experiments.

Gene	Forward Primer	Reverse Primer
B-Tub	TCGTTGGGAGGAGGCACAGGT	GCTGAGTTTGAGGGTACGGAAGCAG
ACS6	CCTGACCGGATTGTTATGAGCGG	TCTATCAAACCTGGATAATAAGGGGTTGG
AT2G21640	CCATGGACAAGATATTCGGGAGCTTCTC	GCTGTCTGAAGAGGAGAGTGAATCTCG
AT3G06530	CGCTGAACCACGAGGTGCTAATG	GCCAAGAATGGAATGTTTCAGCAAGC
ELIP1	CTCACCTTAGGTTTCATAGGAGGAGG	GCAACAGCATCGTTCAACATGC
ELIP2	GCAAGCCTAAGGTGAGTACGAAG	ACCGTCAGAGATCTGAGCAAACACG
NUDT4	CCCGCACAGATCGAGAAAGTTGTC	ACGGTACACATCCGACGACTTGAC
UBQ5	GTAAACGTAGGTGAGTCCA	GACGCTTCATCTCGTCC

2.10 Supplemental materials and methods

2.10.1 Plant Material and Growth Conditions

Arabidopsis (*Arabidopsis thaliana*; ecotype Columbia-0) mutant lines *why1why3* and *why1why3pollb-1* were reported previously [27,43]. Seeds were sown on soil, vernalised for 3 days at 4°C and grown either under normal light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) or low light ($20 \mu\text{mol m}^{-2} \text{s}^{-1}$) at 22°C on a 16 h day/8 h dark cycle. When grown at low light, plants were germinated at $10 \mu\text{mol m}^{-2} \text{s}^{-1}$ for one week prior to their transfer at $20 \mu\text{mol m}^{-2} \text{s}^{-1}$. For drug hypersensitivity experiments, seeds were sterilized and sown on Murashige and Skoog basal media (MS medium, Sigma-Aldrich) supplemented with 1 % sucrose, 0.8 % agar and the indicated amount of drug.

2.10.2 Measurement of PSII Efficiency

For each condition, Fv/Fm measurements were performed on 15-day-old plants treated from days 10 to 15 with 50 μM NOVO, 10 μM CIP or no drug with an Open FluorCam (Photon Systems Instruments) according to the manufacturer's instructions.

2.10.3 Detection of ptDNA Rearrangements

Total DNA was isolated for each plant sample using a cetyl trimethylammonium bromide DNA extraction protocol [249]. For every DNA samples, ptDNA amount was equilibrated with a low-cycle amplification of a *YCF2* DNA fragment. Primers and approaches for rearrangements detection were described previously [29] and PCR products were visualized after migration on GelRed (Biotium) stained agarose gels. Every reaction shown is representative of the eight primer pairs presented in Table I (page 94).

2.10.4 Trypan blue staining

The trypan blue staining of dead cells was conducted as described previously [249] with 4-week-old *Arabidopsis* leaves.

2.10.5 Quantitative RT-PCR Analysis

cDNA synthesis was carried out using a first-strand cDNA synthesis kit (Fermentas) according to manufacturer's instructions. cDNAs were diluted 20-fold prior to qRT-PCR experiments. Primers used for qRT-PCR are listed in Table II (page 94). Every reaction was carried out on biological and technical triplicates and normalised relative to UBQ5. Real-time qPCR reactions were done using the SYBR green master mix (SABiosciences, www.sabiosciences.com) according to the manufacturer's instructions. Melting curves confirmed the amplification of a unique PCR product for each of the qRT-PCR primer pairs used. LightCycler480 (Roche) was used for qRT-PCR experiments and data analyzed using the Light-Cycler480 software version 1.5.

CHAPITRE 3. Caractérisation des relations génétiques entre les gènes *RECA1*, *POLIB*, *WHY1* et *WHY3*

3.1 Mise en contexte

Même si les acteurs assurant le maintien de l'intégrité génomique restent peu connus dans le chloroplaste, certaines protéines clés ont tout de même été identifiées. Ceci inclus l'ADN polymérase POLIB, spécialisée dans la réparation des DSBs, et les protéines Whirly, WHY1 et WHY3, qui répriment des événements de recombinaison illégitime [27,43]. Il a d'ailleurs été démontré par une approche génétique que POLIB, WHY1 et WHY3 travaillent de concert au maintien de la stabilité génomique [43]. Deux homologues de la recombinaise bactérienne RecA sont également retrouvés dans le chloroplaste [207]. De ces deux protéines RecA, plusieurs études ont révélé que RECA1 semble participer activement au maintien de l'intégrité génomique [44,202,209,210]. Aucune étude n'avait toutefois analysé le réseau d'interaction génétique de RECA1 dans le maintien de la stabilité génomique chloroplastique. Une approche génétique a donc été employée avec l'objectif d'évaluer l'existence d'interactions génétiques entre les gènes *RECA1* et *POLIB*, d'une part, et entre *RECA1*, *WHY1* et *WHY3*. Aucune publication des résultats de ce chapitre n'est prévue pour l'instant. Ceux-ci sont présentés pour illustrer la démarche ayant mené à la méthodologie et aux résultats du chapitre quatre. Certains résultats présentés ici seront donc davantage étayés dans le prochain chapitre.

Dans ce chapitre, j'ai généré les lignées doubles mutantes *reca1polla* et *reca1pollb*. J'ai également croisé les lignées mutantes *why1why3* et *reca1* et participé à l'isolement du mutant *why1why3reca1*. J'ai aussi réalisé l'analyse du transcriptome du mutant *why1why3reca1*. De plus, j'ai effectué les expériences d'amplification PCR des réarrangements génomiques, d'évaluation de la sensibilité des lignées mutantes à la CIP, de coloration au bleu de trypan et de caractérisation des paramètres physiologiques du mutant *why1why3reca1*. Thierry Vincent et Étienne Lepage ont contribué à l'isolement du mutant *why1why3reca1*. É. Lepage a contribué à la mesure du paramètre Fv/Fm de *why1why3reca1* et au séquençage de son ARN (ces expériences ont été réalisées simultanément à celles du mutant *why1why3pollb* présentées au chapitre précédent).

3.2 Introduction

Contrairement à la plupart des eucaryotes non-photosynthétiques, les végétaux doivent assurer la stabilité génomique d'un organite supplémentaire, le plastide. Au fil de l'évolution, le transfert vers le noyau de la quasi-totalité des gènes du plastide a mené à un génome extrêmement compact et n'encodant que très peu de gènes. La stabilité génomique de l'ADN du plastide (ADNpt) n'en est pas moins cruciale, les gènes encodés dans ce compartiment étant essentiels au processus de photosynthèse et, conséquemment, à la croissance de la plante. L'intégrité génomique de l'ADNpt est toutefois constamment compromise par les rayons UV-B, les espèces oxygénées réactives (ROS), les bris double-brins de l'ADN (DSBs) et le processus de réplication. Afin de faire face à ces menaces et de maintenir la stabilité génomique, une machinerie protégeant et réparant l'ADN est dédiée au plastide [revu dans 211]. Ces mécanismes semblent d'ailleurs particulièrement efficaces, le taux de mutations de l'ADNpt étant bien plus faible que celui de l'ADN nucléaire [20].

Puisqu'ils peuvent mener à de graves réarrangements chromosomiques s'ils ne sont pas réparés correctement, les DSBs constituent le type de dommage à l'ADN possédant le plus fort potentiel mutagène. Dans le plastide, les principales voies de recombinaison homologue (DSBR et SDSA) sont actives et assurent une réparation dite « conservatrice » des DSBs [191]. Plusieurs enzymes du métabolisme de l'ADN contribuent à ces voies de réparation et empêchent l'accumulation de réarrangements génomiques chloroplastiques. Parmi celles-ci, un homologue de la recombinase bactérienne RecA possède une activité d'échanges de brins d'ADN chez le pois *Pisum sativum* et provoque une hypersensibilité aux agents génotoxiques lorsque muté chez l'algue verte *Chlamydomonas reinhardtii* [208,209]. Chez *Arabidopsis*, deux protéines RecA sont encodées dans le noyau et ciblées au plastide suite à leur expression [207]. Alors que RECA1 n'est retrouvé qu'au plastide, RECA2 est ciblé à la fois au plastide et à la mitochondrie. De manière similaire à *Chlamydomonas*, la mutation *reca1* provoque chez *Arabidopsis* une hypersensibilité à la ciprofloxacine (CIP), un agent génotoxique introduisant des DSBs dans l'ADN du plastide et de la mitochondrie [43,210]. Quant à la mutation *reca2*, celle-ci empêche la germination des plants mutés et,

conséquemment, rend difficile l'étude de ce gène par une approche génétique [213]. Certaines évidences suggèrent néanmoins que la mutation *reca2* affecterait davantage la biogénèse des mitochondries que des chloroplastes [213]. L'ADN polymérase POLIB contribue également au maintien de la stabilité génomique dans le plastide. Le mutant *pol1b* présente en effet une hypersensibilité à la CIP et accumule, lors d'un tel traitement, davantage de DSBs chloroplastiques non-réparés qu'une plante de type sauvage [43]. Puisque cette polymérase semble participer à la réparation des DSBs et que la recombinaison homologue nécessite une étape de synthèse d'ADN, il apparaît possible que POLIB contribue au processus de recombinaison homologue.

Bien qu'ils permettent la réparation des DSBs, les mécanismes de recombinaison doivent néanmoins être finement contrôlés afin d'éviter les réarrangements génomiques. En effet, lors de l'initiation de la recombinaison homologue, la dégradation des brins 5' du DSB génère des extrémités 3' libres simple-brins hautement recombinogéniques et sujettes aux appariements non-spécifiques. Pour cette raison, l'ADN simple-brin (ADNss) est généralement stabilisé par des protéines de liaison. Dans le plastide d'*Arabidopsis*, l'ADNss est retrouvé associé aux protéines Whirly, soit WHY1 et WHY3. La double mutation *why1why3* cause d'ailleurs une hypersensibilité à la CIP et une accumulation de réarrangements génomiques chloroplastiques associés au mécanisme de MMBIR (*Microhomology-Mediated Break-induced Replication*) [27,29]. D'autre part, il a été démontré que la polymérase POLIB travaille de concert avec les protéines Whirly à la réparation adéquate des DSBs [43]. En effet, les individus *why1why3pol1b* présentent un grave phénotype de variégation jaune et une accumulation de réarrangements génomiques bien plus importante que les plants *why1why3* [43,289].

Les évidences citées ci-haut suggèrent toutes la participation des protéines WHY1, WHY3, POLIB et RECA1 à la réparation conservatrice des DSBs chloroplastiques chez *Arabidopsis*. De plus, il a été démontré que la polymérase POLIB travaille conjointement avec les protéines Whirly pour le maintien de l'intégrité génomique chloroplastique [43]. Puisque la recombinaison RECA1 semble également jouer un rôle important dans la réparation des DSBs,

nous postulons que celle-ci travaille également de concert avec les protéines Whirly et POLIB au maintien de la stabilité génomique dans le chloroplaste. Les possibles interactions entre ces gènes ont donc été évaluées par une approche génétique et l'impact de la combinaison de ces mutations sur la stabilité génomique chloroplastique a été caractérisé.

3.3 Résultats et discussion

3.3.1 Caractérisation de la relation génétique entre *RECA1* et *POLIB*

Puisque la recombinaise RECA1 et l'ADN polymérase POLIB semblent toutes deux participer à la réparation de l'ADN, il est possible que ces gènes interagissent génétiquement. Pour tester cette hypothèse, les mutants *reca1* et *pollb* ont été croisés et le double mutant *reca1pollb* a été isolé. Étant donné les rôles de POLIB à la fois au niveau de la réplication et de la réparation de l'ADN, le double mutant *reca1polla* a aussi été généré de manière à pouvoir discriminer ces deux fonctions. En effet, POLIA participe à la réplication de l'ADN mais ne semble pas impliqué dans la réparation [43]. Ainsi, si une interaction génétique est uniquement observée chez le double mutant *reca1pollb*, il apparaîtra probable que celle-ci dépend du rôle de réparation de POLIB.

La combinaison de ces mutations ne cause apparemment pas de phénotype visible, les plants *reca1pola* et *reca1pollb* étant tout-à-fait identiques aux plantes de type sauvage et à chacun des simples mutants (Fig. 32A). Il est toutefois possible que les gènes *RECA1* et *POLIB* ne soient que peu sollicités en absence de stress génotoxique. Afin de tester cette hypothèse, la croissance de chacune de ces mêmes lignées a été comparée en présence de CIP. Tel que démontré précédemment, les mutants *reca1* et *pollb* sont hypersensibles à cet agent génotoxique, au contraire de *polla* qui ne présente aucune sensibilité (Fig. 32A-B) [43,210]. Quant à *reca1polla*, celui-ci présente une hypersensibilité similaire à *reca1*, suggérant qu'aucune interaction génétique n'a lieu entre ces gènes pour la tolérance à la CIP. Au contraire de *reca1polla*, le double mutant *reca1pollb* présente un très fort accroissement de son hypersensibilité comparativement aux mutants *reca1* et *pollb* (Fig. 32A-B). Malgré ce très fort accroissement de l'hypersensibilité à la CIP, il est néanmoins difficile de déterminer si

une interaction génétique existe entre les gènes RECA1 et POLIB. En effet, puisque chacun des simples mutants est sensible à la CIP, il est possible que cette forte hypersensibilité ne résulte que d'un effet additif, et non pas d'une interaction génétique synergique.

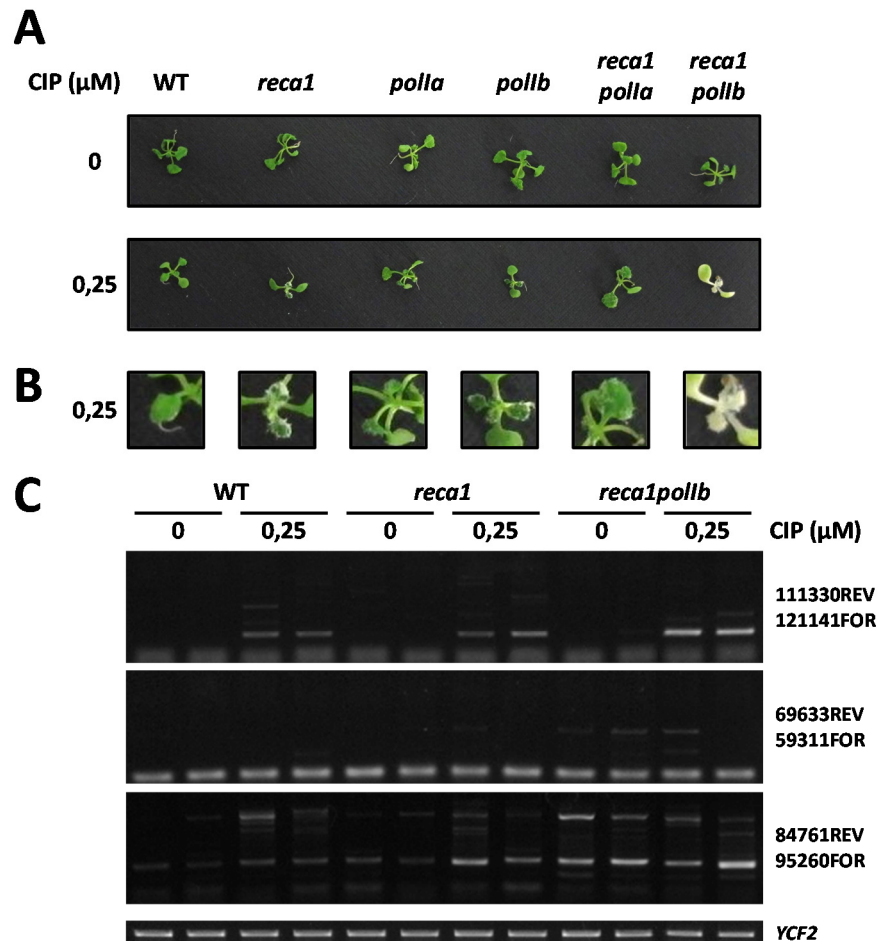


Figure 32. La double mutation *reca1pollb* provoque une grave sensibilité à la CIP, mais ne provoque pas d'accumulation de réarrangements génomiques. (A) Photographies des lignées indiquées âgées de 21 jours lors d'une croissance en absence ou en présence de 0,25 μM CIP. **(B)** Agrandissement des premières vraies feuilles lors d'une croissance en présence de 0,25 μM CIP. **(C)** Réactions PCR représentatives montrant l'abondance de réarrangements génomiques chloroplastiques de type MMBIR chez les lignées indiquées âgées de 21 jours. Les réactions PCR ont été effectuées avec les paires d'amorces indiquées. La quantité d'ADN pour chaque réaction a été équilibrée par rapport à l'amplification d'un court fragment chloroplastique de *YCF2*.

L'hypersensibilité à la CIP des mutants *reca1*, *pollb* et *reca1pollb* suggère que ces deux gènes participent à la réparation des DSBs chloroplastiques. Puisqu'il est probable que ces protéines participent au processus de recombinaison homologue, leur mutation pourrait favoriser des mécanismes de réparation des DSBs alternatifs sujets aux erreurs tels que le MMBIR. Afin de tester cette possibilité, une approche PCR amplifiant spécifiquement les réarrangements génomiques chloroplastiques de type MMBIR et décrite précédemment a été employée [27,29]. Des plants traités à la CIP ont également été inclus dans cette analyse de manière à vérifier si ces voies de réparation alternatives sont davantage sollicitées en conditions de stress génotoxique. Puisqu'il a déjà été démontré que le mutant *pollb* ne semble pas accumuler de réarrangements génomiques, même lorsqu'exposé à la CIP [43], ce mutant n'a pas été inclus dans cette analyse. Aucune augmentation des niveaux de réarrangements génomiques n'est cependant observée chez *reca1* et *reca1pollb*, comparativement aux plants de type sauvage, et ce même en présence de CIP (Fig. 32C).

Chez la levure, une étude récente a démontré que l'accumulation de réarrangements génomiques dépend davantage de l'appariement non-spécifique entre des microhomologies que d'une machinerie de recombinaison homologue non-fonctionnelle. En effet, la mutation de la protéine de liaison à l'ADNss, RPA (*Replication protein A*), semble critique pour l'accumulation de hauts niveaux de ce type de réarrangements [290]. Il apparaît donc probable que, dans le chloroplaste, des protéines de liaison à l'ADNss, telles que WHY1 et WHY3, préviennent l'accumulation de réarrangements génomiques même lorsque la machinerie de recombinaison homologue est fortement amputée. Ceci implique également que, dans de telles conditions, les DSBs resteraient non-réparés plutôt que de mener à des réarrangements génomiques. Une telle éventualité expliquerait d'ailleurs la forte sensibilité à la CIP du mutant *reca1pollb*.

3.3.2 RECA1 est un interacteur génétique des Whirly chloroplastiques

Une défaillance de la recombinaison homologue chloroplastique ne semble pas critique pour la croissance de la plante chez *Arabidopsis* (voir Fig. 32). Il est toutefois possible qu'un mauvais fonctionnement de la recombinaison homologue aggrave l'instabilité génomique

lorsque les voies de réparation sujettes aux erreurs ne sont plus inhibées. Afin de tester cette possibilité, il a été vérifié si une interaction génétique est observée entre les gènes *RECA1*, *WHY1* et *WHY3*, ces deux derniers gènes prévenant l'accumulation de réarrangements génomiques dépendants de microhomologies [27]. Le triple mutant *why1why3reca1* a ainsi été généré et son phénotype a été comparé aux plantes contrôles. La combinaison de ces mutations montre de façon flagrante qu'une interaction génétique a bien lieu entre ces gènes. Le triple mutant *why1why3reca1* présente en effet un important retard de croissance, un phénotype prononcé de variégation blanche et des feuilles déformées (Fig. 33A). Ce phénotype de variégation est en plus associé à une diminution du contenu en chlorophylle dans les tissus de *why1why3reca1* (Fig. 33C). La sévérité de ce phénotype suggère que les mutations *why1why3reca1* ont de graves effets sur le développement de la plante. Pour vérifier si ce phénotype est associé à un taux accru de mort cellulaire, les cellules mortes des feuilles de *why1why3reca1* ont été teintées avec le colorant bleu de trypan. Comparativement aux plantes de type sauvage, aucune augmentation du taux de mort cellulaire n'est toutefois observée chez *why1why3reca1* (Fig. 33B). Les mutations *why1why3reca1* affectent donc le développement normal des tissus, mais n'affecte pas leur viabilité.

Puisque les mutants *reca1* et *why1why3* sont tous deux hypersensibles à la CIP, il est possible que la synergie de ces mutations se perçoive également par la sensibilité aux stress génotoxiques affectant le chloroplaste. Pour tester cette hypothèse, la croissance de *why1why3reca1* en présence de CIP a été comparée aux plantes simples mutantes. Même s'ils sont hypersensibles à la CIP, les mutants *reca1* et *why1why3* réussissent à développer des secteurs verts sur leurs feuilles affectées (Fig. 33D). Au contraire de ces mutants, les feuilles affectées par la CIP chez *why1why3reca1* ne développent jamais de secteurs verts (seuls les cotylédons sont verts). Cette forte hypersensibilité et le phénotype du mutant *why1why3reca1* suggèrent donc qu'une importante interaction génétique a lieu entre les gènes *RECA1*, *WHY1* et *WHY3* et ce, probablement, au niveau du maintien de l'intégrité génomique chloroplastique. Tout comme pour le mutant *reca1pollb*, il est néanmoins

impossible d'exclure complètement que cette forte hypersensibilité à la CIP ne résulte que d'un effet additif.

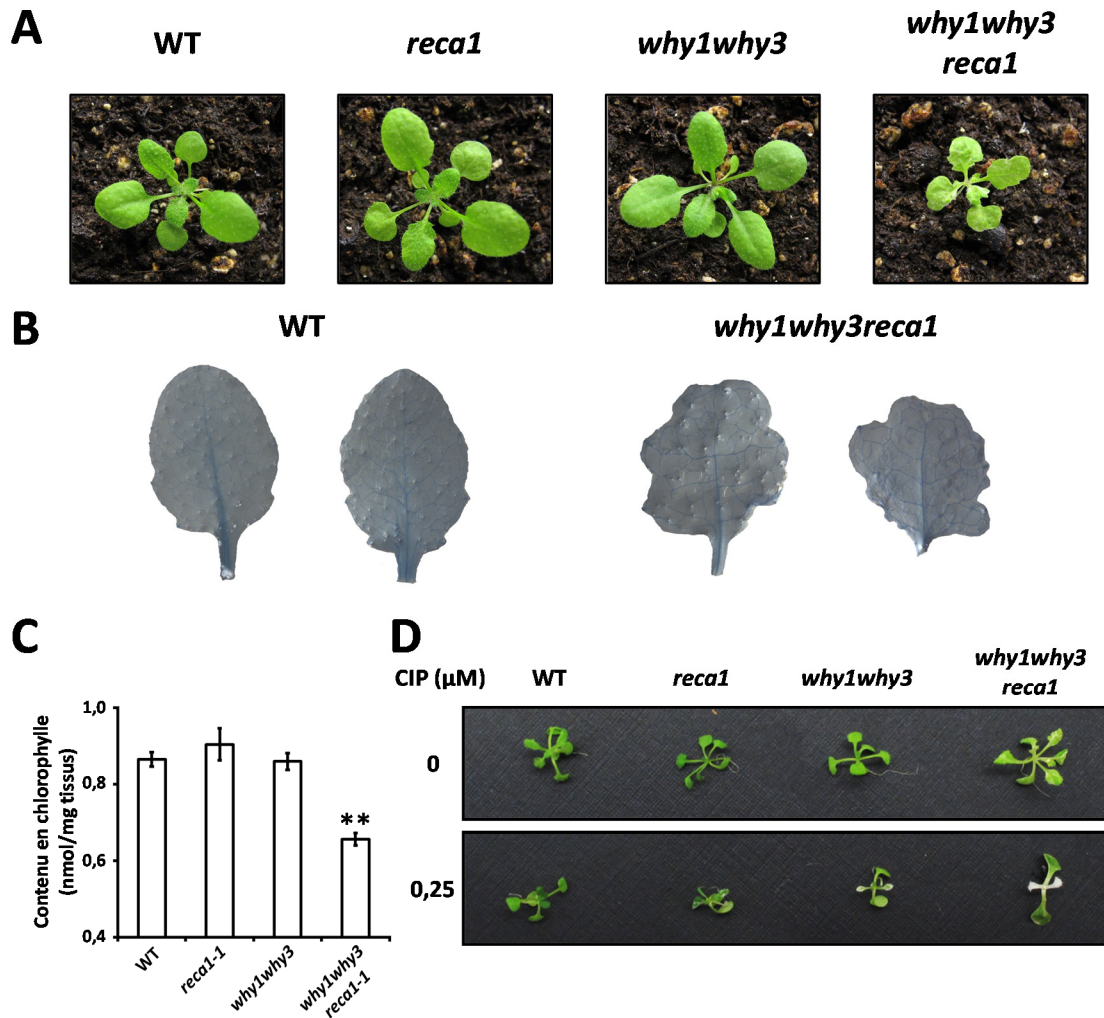


Figure 33. Le gène *RECA1* est un interacteur génétique des Whirly chloroplastiques. (A) Photographies d'une plante de type sauvage (WT) et des lignées indiquées âgées de 21 jours. (B) Photographies de feuilles WT et *why1why3reca1* âgées de 28 jours et teintées au bleu de trypan. (C) Contenu en chlorophylle de chacune des lignées indiquées âgées de 21 jours. Les barres d'erreur représentent l'erreur standard de cinq mesures effectuées sur des échantillons biologiques distincts. Deux astérisques représentent une différence significative avec le WT d'un test de t à une valeur $P \leq 0,01$. (D) Photographies des lignées indiquées âgées de 21 jours lors d'une croissance en absence ou en présence de 0,25 μM CIP.

3.3.3 Les gènes *RECA1*, *WHY1* et *WHY3* travaillent de concert au maintien de la stabilité génomique chloroplastique

Il a été démontré précédemment que la combinaison de la mutation *pollb* et des mutations *why1why3* a un effet synergique sur l'accumulation de réarrangements génomiques dépendants de microhomologies [43]. Puisque *RECA1* semble impliqué dans les voies de réparation de l'ADNpt, il est donc possible qu'en plus du phénotype variégation blanche, la synergie des mutations *why1why3reca1* se perçoive au niveau de l'instabilité génomique chloroplastique. Pour tester cette hypothèse, la même approche PCR a été employée. Tel que démontré auparavant, les mutations *reca1* et *pollb* ne causent aucune accumulation notable de réarrangements génomiques (Fig. 34A). Quant aux triples mutants *why1why3pollb* et *why1why3reca1*, tous deux présentent des niveaux bien plus élevés de réarrangements génomiques que le mutant *why1why3* (Fig. 34A). Cette synergie dans l'accumulation de réarrangements génomiques chez ces deux triples mutants suggère qu'une interaction génétique existe bien entre les protéines Whirly et la machinerie de réparation de l'ADN, probablement par des voies de recombinaison homologue. Ce résultat suggère également que les gènes *RECA1*, *POLIB*, *WHY1* et *WHY3* participent tous au maintien de la stabilité génomique dans le chloroplaste.

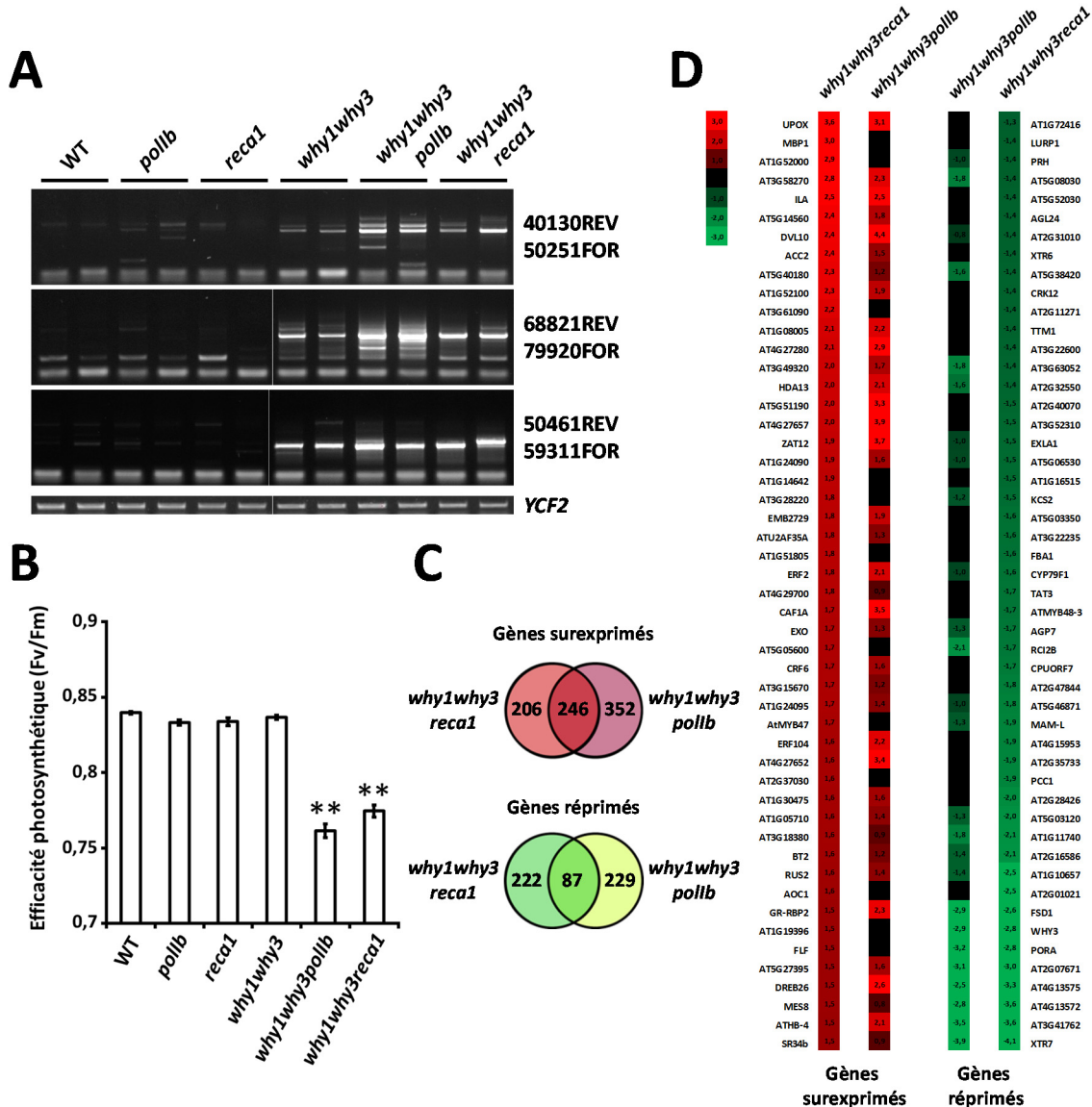


Figure 34. Les mutations *why1why3reca1* et *why1why3pollb* mènent à une forte accumulation de réarrangements génomiques chloroplastiques et à une signalisation rétrograde du chloroplaste au noyau. **(A)** Réactions PCR représentatives montrant l'abondance de réarrangements génomiques chloroplastiques de type MMBIR chez les lignées indiquées âgées de 28 jours. Les réactions PCR ont été effectuées avec les paires d'amorces indiquées. La quantité d'ADN pour chaque réaction a été équilibrée par rapport à l'amplification d'un court fragment chloroplastique de *YCF2*. **(B)** Efficacité photosynthétique (Fv/Fm) des lignées indiquées. Les barres d'erreur représentent l'erreur standard d'au moins neuf mesures faites sur des plants distincts. Deux astérisques représentent une différence significative avec le WT d'un test de t à une valeur $P \leq 0,01$. **(C)** Diagramme de Venn représentant le nombre de gènes surexprimés ou

réprimés d'au moins 1,75 fois par rapport aux plantes de type sauvage chez *why1why3reca1* et *why1why3pollb* et communs à chacune des deux lignées. **(D)** Heatmap des 50 gènes les plus surexprimés et les plus réprimés chez *why1why3reca1* et comparaison à l'expression de chacun de ces gènes chez *why1why3pollb*.

3.3.4 Les mutants *why1why3pollb* et *why1why3reca1* possèdent des patrons similaires de reprogrammation génique nucléaire

Étonnamment, malgré une augmentation similaire de l'instabilité génomique dépendante de microhomologies chez *why1why3pollb* et *why1why3reca1*, ces deux triples mutants présentent tout de même des phénotypes bien distincts. Chez *why1why3pollb*, cette forte instabilité génomique est associée à un déséquilibre de la chaîne de transport des électrons [289]. Afin de vérifier si la forte instabilité génomique observée chez *why1why3reca1* mène elle aussi à un tel déséquilibre, la capacité photosynthétique de cette lignée a été déterminée par la mesure du paramètre Fv/Fm (Fig. 34B). Cette mesure révèle que, comme chez *why1why3pollb*, les réarrangements génomiques observés chez *why1why3reca1* corréleront avec une diminution de la valeur Fv/Fm et, conséquemment, avec un déséquilibre de la chaîne de transport des électrons.

Chez *why1why3pollb*, ce déséquilibre de la chaîne de transport des électrons est associé à un déséquilibre rédox qui, à son tour, mène à une reprogrammation génique nucléaire [289]. De plus, cette reprogrammation génique corrèle avec l'apparition du phénotype de variégation jaune de cette lignée, tous deux étant complémentés en conditions de faible déséquilibre rédox. Il est ainsi possible que la différence de phénotype entre les deux triples mutants provienne de patrons distincts d'expression génique nucléaire. Pour tester cette hypothèse, une approche de séquençage par RNAseq a été employée pour caractériser le transcriptome nucléaire de *why1why3reca1*. Comparativement aux plantes de type sauvage, 452 et 598 gènes sont surexprimés et réprimés d'au moins 1,75 fois, respectivement, chez *why1why3reca1*.

D'autre part, la comparaison des transcriptomes de *why1why3reca1* et *why1why3pollb* révèle une très grande similarité entre leurs patrons de reprogrammation génique (Fig. 34C-D). En effet, environ 54 % des gènes surexprimés plus de 1,75 fois chez *why1why3reca1* sont également surexprimés chez *why1why3pollb*. Quant aux gènes réprimés d'au moins 1,75 fois chez *why1why3reca1*, 28 % le sont aussi chez *why1why3pollb*. De manière plus générale, il a été montré que la reprogrammation génétique nucléaire du mutant *why1why3pollb* ressemble fortement à celles observées en conditions de stress abiotiques et oxydatifs [289]. La très grande similarité des patrons transcriptomiques des deux triples mutants suggère donc que la reprogrammation génétique observée chez *why1why3reca1* est également associée à un déséquilibre rédox chloroplastique. De plus, tout comme chez *why1why3pollb*, plusieurs marqueurs associés au stress oxydatif sont surexprimés chez *why1why3reca1*. Ceux-ci incluent les facteurs de transcription ZAT10 et ZAT12, associés au stress oxydatif et aux hautes intensités lumineuses [111,112], l'ARNm déadénylase CAF1A impliquée dans la réponse aux stress abiotiques [259] et le marqueur de stress oxydatif *UPOX* [106,265]. Il apparaît donc probable que les réarrangements génomiques causés par les mutations *why1why3pollb* et *why1why3reca1* soient à l'origine de cette reprogrammation et de la très grande similarité de leurs patrons transcriptomiques. Ainsi, bien que ces deux lignées présentent des phénotypes différents, les mutations *why1why3pollb* et *why1why3reca1* semblent toutes deux provoquer une déstabilisation de l'ADNpt, un déséquilibre photosynthétique et une reprogrammation génétique associée au stress oxydatif.

3.3.5 Le phénotype de variégation blanche de *why1why3reca1* n'est pas affecté par l'intensité lumineuse

Chez *why1why3pollb*, l'apparition du phénotype de variégation jaune peut être modulée par l'intensité lumineuse lors de la croissance de la plante. Puisque l'intensité lumineuse influence directement la production de ROS par la PET, il semble probable que l'apparition de la variégation jaune soit la conséquence du déséquilibre rédox chloroplastique observé chez *why1why3pollb* [289]. Afin de vérifier si le phénotype de variégation blanche du mutant *why1why3reca1* peut aussi être modulé par la concentration de ROS chloroplastiques, la

croissance de cette lignée a été comparée dans des conditions de faible et de haute intensité lumineuse. Toutefois, contrairement à *why1why3pollb*, le phénotype de variégation blanche n'est ni accentué à haute intensité lumineuse, ni complémenté à faible intensité (Fig. 35). Le phénotype de variégation blanche du triple mutant *why1why3reca1* ne semble donc pas être associé à un déséquilibre rédox chloroplastique.

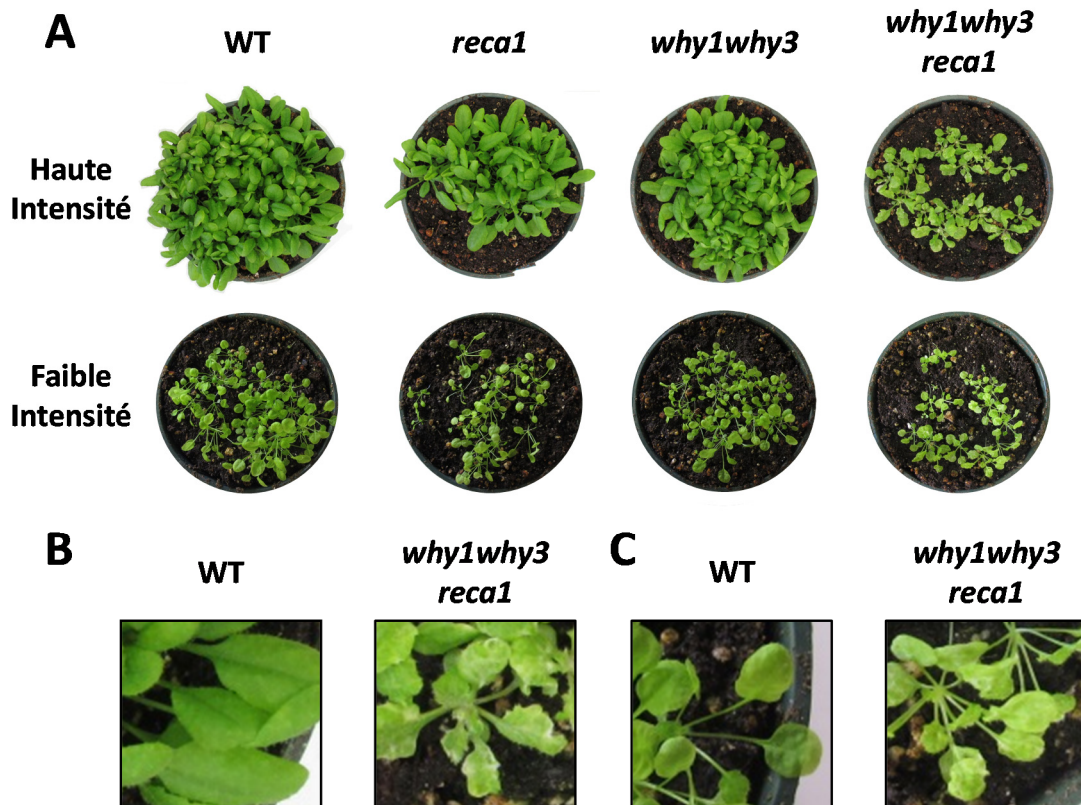


Figure 35. Le phénotype de variégation blanche n'est pas affecté par l'intensité lumineuse chez *why1why3reca1*. (A) Photographies de chacune des lignées indiquées lors d'une croissance à haute ou à faible intensité lumineuse pour 28 et 60 jours, respectivement. (B) Agrandissement de tissus lors d'une croissance à haute intensité lumineuse. (C) Agrandissement de tissus lors d'une croissance à faible intensité lumineuse.

3.4 Conclusion

L'étude des interactions génétiques existant entre les gènes *RECA1*, *POLIB*, *WHY1* et *WHY3* suggère un important rôle pour la protéine RECA1 dans le maintien de la stabilité génomique chloroplastique. Les résultats obtenus jusqu'ici suggèrent en effet que RECA1 travaille de concert avec POLIB et avec les protéines Whirly à la réparation des DSBs, les mutants *reca1pollb* et *why1why3reca1* étant tous deux extrêmement sensibles à la CIP. Il est néanmoins difficile de déterminer si RECA1 et POLIB effectuent leurs fonctions au sein d'une même voie de réparation de l'ADN, la forte hypersensibilité à la CIP du mutant *reca1pollb* pouvant provenir autant d'un effet additif des mutations que d'un effet synergique. Quant au triple mutant *why1why3reca1*, l'apparition du sévère phénotype de variégation blanche et l'effet synergique des mutations sur l'accumulation de réarrangements génomiques de type MMBIR suggèrent qu'une interaction génétique synergique existe entre les gènes *RECA1*, *WHY1* et *WHY3*. Les interactions génétiques synergiques sont habituellement observées lorsque des mutations affectent des voies distinctes, mais possédant un nœud ou une étape commune. Cette interaction chez *why1why3reca1* suggère que la répression des événements de MMBIR par les protéines Whirly et la recombinaison homologue dépendante de RECA1 constituent des voies distinctes, mais travaillant conjointement à la réparation conservatrice des DSBs. Il semble donc probable que l'inactivation simultanée de ces deux voies favoriserait de façon synergique les voies de réparation sujettes aux erreurs, telles que le MMBIR.

Il est surprenant de constater que malgré une accumulation similaire de réarrangements génomiques chloroplastiques, les deux triples mutants *why1why3pollb* et *why1why3reca1* ont tout de même des phénotypes distincts. L'incapacité à moduler le phénotype de variégation blanche de *why1why3reca1* par l'intensité lumineuse suggère que ces mutations ont d'autres conséquences en plus de celles détectées jusqu'ici. Chez la bactérie, en plus de son rôle dans la réparation des DSBs, la protéine RecA participe également au redémarrage des fourches de réplication bloquées. Quant aux protéines de liaison à l'ADNss nommées SSBs (*Single-strand DNA-binding proteins*), celles-ci participent aussi à la fois aux processus

de réplication et de réparation des DSBs. De tels exemples soulèvent donc la possibilité que les mutations *why1why3reca1* affectent d'autres aspects du métabolisme de l'ADNpt qui pourraient être responsables de l'apparition de ce très sévère phénotype de variégation blanche.

3.5 Méthodes

Les lignées mutantes étudiées dans ce chapitre sont décrites au chapitre 4. Les conditions de croissance sur terre en conditions d'intensité lumineuse normale, haute et faible sont décrites au chapitre 2. Les expériences d'amplification PCR des réarrangements génomiques de type MMBIR, de coloration au bleu de trypan, de mesure des niveaux de chlorophylle, de mesure de l'efficacité photosynthétique (Fv/Fm) et de séquençage du transcriptome nucléaire sont décrites au chapitre 2. Pour la croissance des plants en présence ou en absence de CIP, les graines ont été stérilisées et plantées sur milieu Murashige and Skoog (MS; Sigma-Aldrich) additionné de 1% sucrose (w/v), 0,8% agar (w/v), 0,05% MES pH 5,7 (w/v) et la quantité indiquée de CIP.

CHAPITRE 4. Organelle DNA rearrangement mapping reveals U-turn-like inversions as a major source of genomic instability in *Arabidopsis* and humans

**Organelle DNA rearrangement mapping
reveals U-turn-like inversions as a major
source of genomic instability in *Arabidopsis*
and humans**

Publié dans Genome Research, mai 2015

Vol. 25, numéro 5, pages 645–654

Éric Zampini*, Étienne Lepage*, Samuel Tremblay-
Belzile*, Sébastien Truche, et Normand Brisson

*Ces auteurs ont contribué également à l'article

4.1 Mise en contexte

Par une approche génétique, il avait été observé que les protéines Whirly travaillent de concert avec les protéines POLIB et RECA1 au maintien de l'intégrité génomique dans le chloroplaste. En effet, la combinaison de ces mutations dans les plants *why1why3pollb* et *why1why3reca1* provoque une importante accumulation de réarrangements génomiques associés au processus de MMBIR. De plus, chez ces deux lignées, cette déstabilisation de l'ADNpt mène à un déséquilibre photosynthétique et à une reprogrammation génique nucléaire. Pourtant, les phénotypes particuliers et distincts de chacune de ces lignées ne reflètent pas cette apparente similarité des conséquences moléculaires. Puisque le phénotype de variégation jaune du mutant *why1why3pollb* avait été associé à un déséquilibre rédox chloroplastique, il semblait probable que les mutations *why1why3reca1* aient d'autres conséquences que celles détectées jusqu'ici. Les gènes *WHY1*, *WHY3* et *RECA1* participant tous au métabolisme de l'ADNpt, il apparaissait plausible que les mutations de ces gènes provoquent l'accumulation de réarrangements génomiques particuliers et passés inaperçus par l'utilisation de l'approche PCR utilisée pour détecter les réarrangements de type MMBIR. Afin de caractériser le patron d'instabilité génomique chloroplastique de manière globale et non biaisée, l'ADNpt de chacune de ces lignées a été séquencé par séquençage Illumina et les réarrangements génomiques ont été analysés par une approche bio-informatique.

Dans cet article, j'ai effectué les expériences préliminaires d'optimisation du séquençage de nouvelle génération et l'extraction des échantillons d'ADN analysés dans ce manuscrit. J'ai également participé au développement de l'approche bio-informatique, à l'analyse des patrons d'instabilité génomique et à la caractérisation des types de réarrangements génomiques séquencés. De plus, j'ai réalisé les expériences de PCR quantitatif et de PFGE présentées dans le manuscrit. É. Lepage et S. Tremblay-Belzile ont participé au développement de l'approche bio-informatique, ainsi qu'à l'analyse des résultats. L'interprétation des résultats, l'écriture et la révision du manuscrit ont été effectuées

conjointement par É. Lepage, S. Tremblay-Belzile, N. Brisson et moi-même. S. Truche a également participé à l'interprétation des résultats.

4.2 Abstract

Failure to maintain organelle genome stability has been linked to numerous phenotypes, including variegation and cytosolic male sterility (CMS) in plants, as well as cancer and neurodegenerative diseases in mammals. Here we describe a next-generation sequencing approach that precisely maps and characterizes organelle DNA rearrangements in a single genome-wide experiment. In addition to displaying global portraits of genomic instability, it surprisingly unveiled an abundance of short-range rearrangements in *Arabidopsis thaliana* and human organelles. Among these, short-range U-turn-like inversions reach 25% of total rearrangements in wild-type *Arabidopsis* plastids and 60% in human mitochondria. Furthermore, we show that replication stress correlates with the accumulation of this type of rearrangements, suggesting that U-turn-like rearrangements could be the outcome of a replication-dependent mechanism. We also show that U-turn-like rearrangements are mostly generated using microhomologies and are repressed in plastids by Whirly proteins WHY1 and WHY3. A synergistic interaction is also observed between the genes for the plastid DNA recombinase *RECA1* and those encoding plastid Whirly proteins, and the triple mutant *why1why3reca1* accumulates almost 60 times the WT levels of U-turn-like rearrangements. We thus propose that the process leading to U-turn-like rearrangements may constitute a RecA-independent mechanism to restart stalled forks. Our results reveal that short-range rearrangements, and especially U-turn-like rearrangements, are a major factor of genomic instability in organelles and raise the question of whether they could have been underestimated in diseases associated with mitochondrial dysfunction.

4.3 Introduction

The endosymbiosis events at the origin of mitochondria and plastids have been a key step in the appearance of eukaryotic cells. During evolution, most of the genes present in the

genome of the endosymbionts have been lost or transferred to the nucleus, resulting in organelles with small, dense genomes that almost exclusively encode essential genes. There is therefore a strong selective pressure for the maintenance of genome stability in organelles [20]. In plants, *Arabidopsis thaliana* has emerged as a useful model to study organelle genome rearrangements. Several actors involved in the maintenance of *Arabidopsis* plastid and mitochondrion genomes have been identified, among which are the Whirly proteins, a family of single-stranded DNA-binding proteins that guard organelles against genomic rearrangements [27,29]. It has been proposed that these proteins stabilize single-stranded DNA and guide it through conservative repair mechanisms such as homologous recombination [27,29]. Whirly proteins are found in all plant organelles and, in *Arabidopsis*, WHY1 and WHY3 are targeted to the plastids while WHY2 is targeted to the mitochondria [217]. The RecA family of DNA-binding proteins also maintains organelle genome stability through their central role in homologous recombination. In *Arabidopsis*, three RecA proteins are found in organelles, with RECA1 targeted to the chloroplast, RECA3 to the mitochondrion, and RECA2 targeted to both organelles [207]. These proteins were shown to be involved in DNA double-strand break (DSB) repair and, in plastids, to maintain the structure of the genome [210,213]. Prokaryote RecA proteins were also shown to be essential for fork reversal, a mechanism that allows the accurate restart of paused replication forks, thereby promoting fork progression in conditions of replication stress [197,291,292].

Replication represents a major challenge to plastid and mitochondrion genome stability. For example, it was shown that the mutation of the type-I polymerase *POLIB* in *Arabidopsis* causes replication stress at early developmental stages and increases the amount of DSBs upon genotoxic stress treatment [43]. Interestingly, mutation of the mammalian mitochondrial DNA polymerase gamma has also been linked to replication stress and DSBs [293,294]. Replication-dependent DSBs are known to arise when a fork collapses or encounters a nick in the matrix DNA [152]. These DSBs can subsequently be repaired by homologous recombination or by error-prone mechanisms such as microhomology-mediated recombination (MHMR) and non-homologous end-joining (NHEJ) [200,295,296]. Replication-

dependent genomic instability is however not solely induced by DSB repair mechanisms [152]. Indeed, stalled forks have recently been shown to produce DSB-independent fusions of nearby inverted-repeats, which lead to the formation of palindromic chromosomes [297,298]. The mechanisms by which these fusions take place remain a subject of debate, and three distinct possibilities have been proposed: faulty template switching, tandem inversion duplications, and replication U-turns [297-301].

To date, our understanding of how organisms deal with replication-associated genomic instability has mostly been obtained using reporter systems. A drawback of these systems is that they are impractical for the study of organelle genomes. To overcome this limitation, we developed a next-generation sequencing approach which allows the characterization of organelle DNA instability at a genome-wide level while providing information about the mechanism underlying each rearrangement formation.

4.4 Results

4.4.1 Short-range rearrangements are abundant in organelle genomes

We first evaluated overall plastid DNA (ptDNA) rearrangements in the *Arabidopsis thaliana* ecotype Col-0. To obtain a global and quantitative portrait of ptDNA instability, a heat map representation was employed in which all rearrangements are reported at the intersection of the two genomic coordinates that correspond to each side of the rearrangement junction. This analysis shows that many rearrangements occur apparently randomly in this plastid genome (Fig. 36A and Supplemental Table S1, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>). It also reveals an over-representation of short-range rearrangements (<1,000 bp), as indicated by the higher intensity of the heat map diagonal (Fig. 36A,D). To validate our approach, the same analysis was repeated with the *Arabidopsis thaliana* ecotypes Ts-1 (Tossa de Mar) and Ws-2 (Wassilewskija) [302]. Results indicate that the overall pattern and relative level of plastid DNA rearrangements are strikingly similar between ecotypes (Fig. 42, page 143, and Supplemental Tables S2-3, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>).

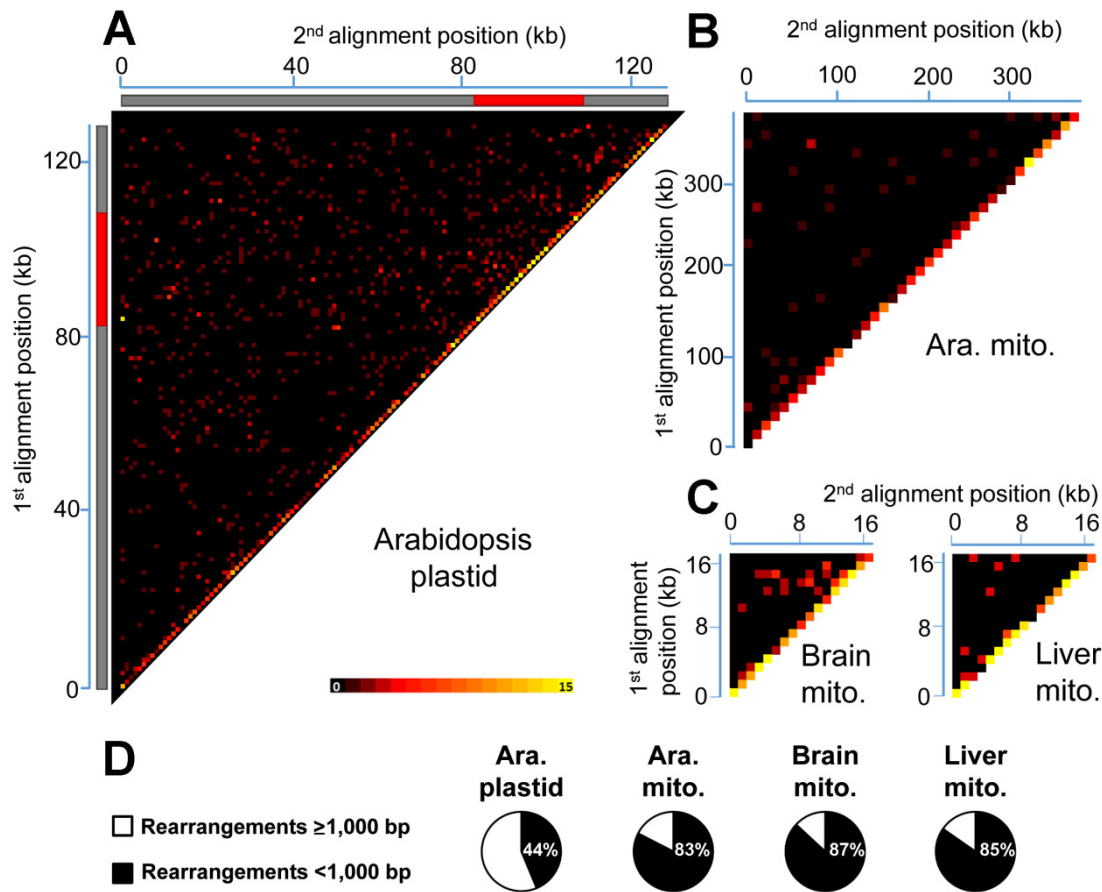


Figure 36. Global portrait of organelle genome rearrangements in *Arabidopsis* and human. Heat maps depict each rearrangement as the intersection of the two genomic positions corresponding to the nucleotide on each side of the junction. **(A)** Rearrangement breakpoint positions of *Arabidopsis thaliana* (Col-0) plastids. Each tile represents a region spanning 1 kb along each axis. Tile intensity represents the number of rearrangements per 10,000 plastid genomes. All rearrangements mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. The plastid large-single copy region (LSC), the first IR, and the small-single copy region (SSC) are depicted as a long grey bar, a red bar and a short grey bar, respectively. **(B)** Rearrangement breakpoint positions of wild-type *Arabidopsis thaliana* (Col-0) mitochondria. Each tile represents a region spanning 10 kb along each axis. Tile intensity represents the number of rearrangements per 100 mitochondrion genomes. **(C)** Rearrangement breakpoint positions of a representative sample for human brain (ERR419275) and liver (ERR419281) mitochondria. Each tile represents a region spanning 1 kb along each axis. Tile intensity represents the number of rearrangements per 10,000 mitochondrion genomes. **(D)** Proportion of short-range (breakpoint positions less than 1,000 bp

apart) and long-range (breakpoint positions at least 1,000 bp apart) rearrangements of *Arabidopsis thaliana* (Col-0) plastid, mitochondrion and of the mean of the four samples for brain and liver mitochondria. Ara.: *Arabidopsis*, mito.: mitochondrion.

To verify if mitochondria exhibit a similar pattern of genomic instability, we generated a heat map for *Arabidopsis thaliana* (Col-0) mitochondrial genome rearrangements. Interestingly, the vast majority of rearrangements also occur at short range in this organelle (Fig. 36B,D and Supplemental Table S4, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>). To our knowledge, this is the first time that such a high level of short-range rearrangements is reported in organelle genomes, which raises the question of whether these could also be abundant in animal mitochondria. To test this hypothesis, we subjected publicly available datasets from human brain and liver to the same analysis. This revealed a similar pattern of short-range genomic instability, reaching 86% and 84% of total rearrangements in brain and liver cells, respectively (Fig. 36C-D and Supplemental Tables S5-12, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>). In contrast, analysis of genomic instability in *E. coli* reveals a much less striking prominence of short-range rearrangements (Supplemental Tables S13-16, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>), indicating that the detection of high levels of these rearrangements in organelles does not occur systematically in our approach. Globally, these results indicate that short-range rearrangements are a major factor of genomic instability in both plant and animal organelles.

4.4.2 Microhomology and non-microhomology repair happen at similar rates in wild-type organelles

To get insights into the mechanisms involved in the formation of the DNA rearrangements in organelles, we further analyzed the reads corresponding to rearranged genome molecules. Because these reads are always composed of two alignments mapping distinct regions of the genome, two types of DNA rearrangements can easily be discriminated: those that possess a microhomology at their junction, most likely formed by MHMR, and those without microhomology, reminiscent of NHEJ repair (Fig. 37A). This analysis revealed that in wild-type Col-0 (WT) plastids and mitochondria, 55% and 64% respectively of genome rearrangements

detected arose from microhomology-dependent pathways with microhomologies of 5 bp or more (Fig. 37B). Similarly, rearrangements dependent on microhomologies represent 64% and 67% of total genomic instability in human brain and liver mitochondria, respectively (Fig. 37B).

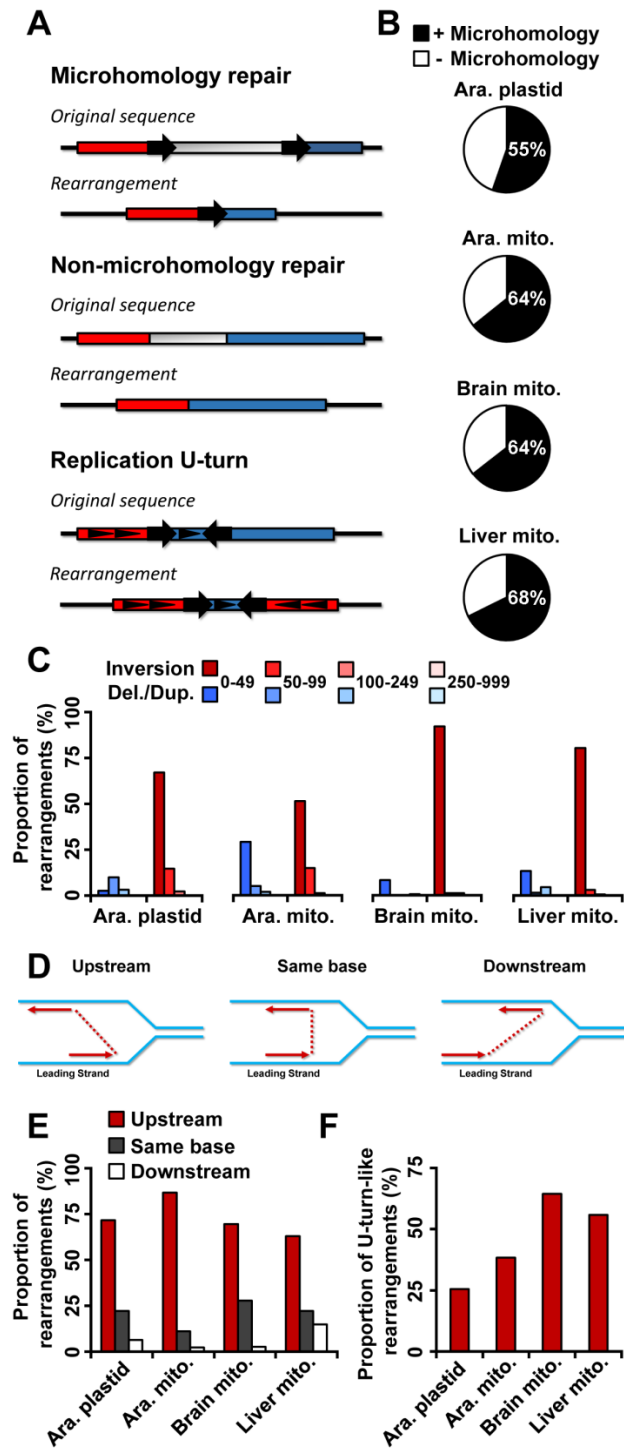


Figure 37. Analysis of organelle genome rearrangements in *Arabidopsis* and human. (A) Depiction of the types of rearrangement junctions observed by next-generation sequencing. Large arrows represent microhomologies and their orientation. Small triangles represent relative strand directions. (B) Proportion of

rearrangements identified displaying a microhomology (≥ 5 bp) at their junction in each organelle. The mean of the four samples for brain and liver mitochondria is presented. **(C)** Proportion of short-range deletions/duplications and inversions ($<1,000$ bp) displaying a gap of the indicated length. Y axis represents the percentage of each type of rearrangement relative to its class size (gap length). A representative sample is shown for human brain (ERR419275) and liver (ERR419281) mitochondria. **(D)** Schematic representation of short-range inversions displaying a junction upstream, at the same base, or downstream on the opposite strand. The matrix strands are shown in blue, and the nascent strands in red with the junction gap shown as a dotted line. **(E)** Proportion of short-range inversions (<50 bp) displaying a junction upstream, at the same base, or downstream on the opposite strand in each organelle genome. A representative sample is shown for human brain (ERR419275) and liver (ERR419281) mitochondria. **(F)** Proportion of total rearrangements corresponding to short-range inversions (<50 bp, U-turn-like rearrangements) in each organelle genome. The mean of the four samples for brain and liver mitochondria is presented. Ara.: *Arabidopsis*, mito.: mitochondrion, del.: deletion, dup.: duplication.

4.4.3 U-turn-like rearrangements are ubiquitous among organelle short-range rearrangements

The fact that short-range rearrangements represent a large part of the total junctions identified in organelle genomes suggests that these rearrangements could arise through a specific mechanism taking place at nearby microhomologies. Analysis of rearrangement junctions at the base pair level offers the key advantage of allowing the characterization of these mechanisms. We therefore measured the proportion of rearrangements involving inversions among the short-range rearrangements ($<1,000$ bp) according to their gap length in comparison to other types of rearrangements (duplications and deletions) in *Arabidopsis* and human organelles. This revealed that inversions occurring over distances smaller than 50 bp are largely overrepresented (Fig. 37C). This type of genomic instability seems to be of particular importance in organelles since it represents approximately 26% and 38% of total rearrangements in *Arabidopsis* plastids and mitochondria, respectively (Fig. 37F). This proportion is even higher in human mitochondria, reaching 64% and 56% of the total brain and liver genomic instability, respectively (Fig. 37F). A closer look at the junctions of these inversions indicates that they mainly occur upstream of the junction breakpoint, on the

opposite strand (Fig. 37D-E). Interestingly, although often imperfect, approximately 85% of these short-range inversions present a microhomology at their junction in plastids. The overall characteristics of the DNA junctions observed in these short range inversions can best be explained by a model in which replication executes a U-turn upstream on the opposite strand (Fig. 37D). It also suggests that the inversions take place on the opposite strand between the 5' end of the lagging strand and the site where the DNA is unwound in the replication fork.

4.4.4 Whirly proteins protect the plastid genome from MHMR but not from NHEJ

The organelle DNA polymerase POLIB and Whirly proteins have previously been shown to protect plastid genomes against long-range deletions and duplications mediated by microhomologies. To verify whether these proteins do also offer a protection against short-range DNA rearrangements, we further characterized the *Arabidopsis* mutant lines *why1why3* and *pollb* by sequencing. We also included in our analysis the mutant line *reca1-1* (hereafter called *reca1*), which is mutated for plastid *RECA1* (Fig. 43, page 144). While only a small increase in ptDNA rearrangements is observed in *pollb* and *reca1* as compared to the Col-0 parent line, a much higher level of genomic instability is observed in the *why1why3* line (Fig. 38A-B and Supplemental Tables S17-19, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>). Interestingly, the heat map diagonal region of all three mutant lines also shows a much higher intensity than the rest of the genome (Fig. 38A). It thus seems that short-range rearrangements also constitute an important part of the total ptDNA instability in these lines and that their occurrence is increased when genes involved in DNA metabolism are mutated. To get insights into the mechanisms involved in the formation of these rearrangements, we determined the prevalence of microhomology usage for each line. While approximately 0.12 DNA rearrangements per genome arose from microhomologies in WT plants, this increased to 0.8 microhomology-dependent rearrangements per genome in *why1why3* (Fig. 38B). This shift toward microhomology usage in *why1why3* is also confirmed by the analysis of the lengths of the microhomologies leading to rearrangements. Indeed, while microhomology-

mediated rearrangements in WT, *reca1* and *pollb* plants are in large part produced by microhomologies of 5 to 9 bp, those generated from microhomologies of 10 to 14 bp are the most prevalent in the Whirly mutant (Fig. 38C). In contrast, the level of DNA rearrangements generated independently of microhomologies is similar in all three lines (Fig. 38B), suggesting that Whirly proteins mainly suppress the appearance of microhomology-dependent rearrangements, and do not affect microhomology-independent pathways. The *pollb*, *reca1* and *why1why3* mutations however have little effect on the accumulation of U-turn-like rearrangements (inversions occurring over distances smaller than 50 bp), which only slightly increase in all three mutant lines (Fig. 38D).

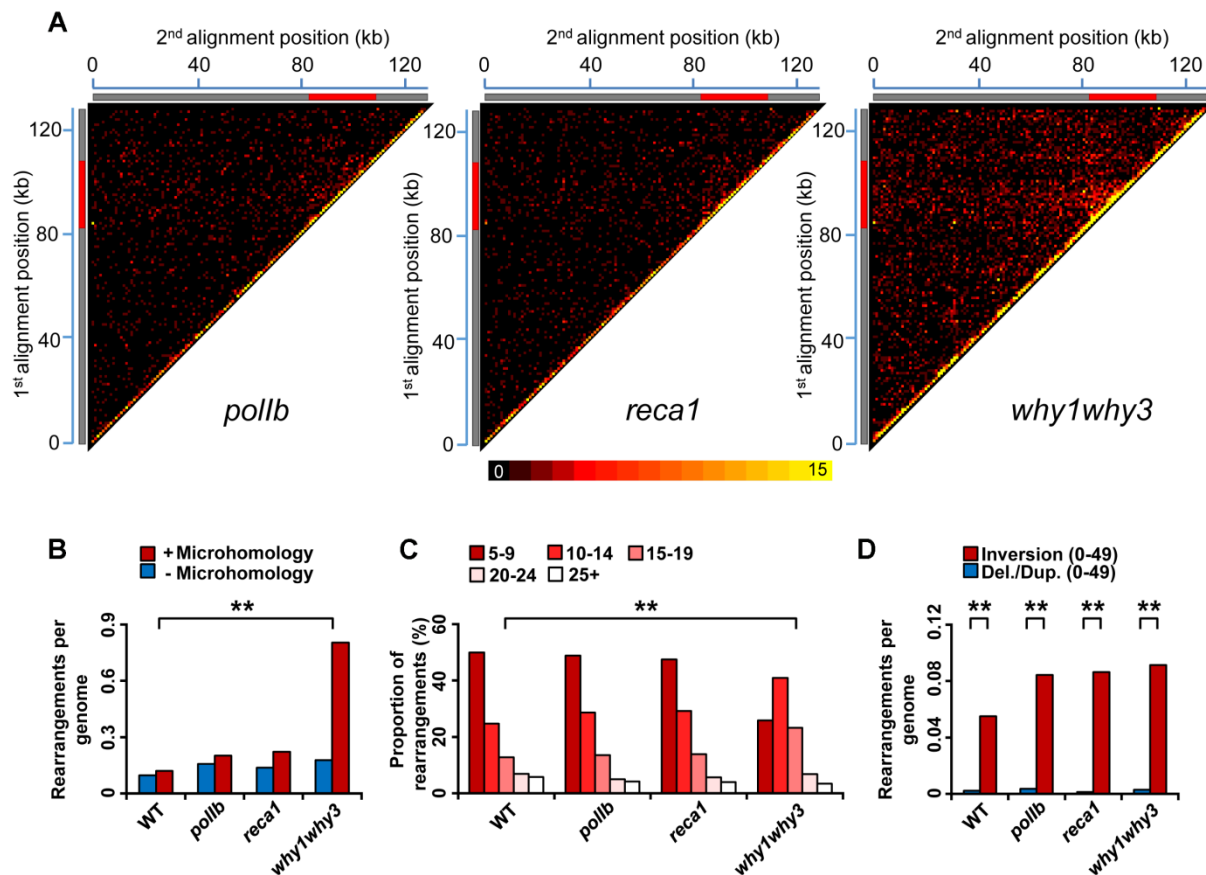


Figure 38. Global portrait of plastid genome rearrangements in *Arabidopsis* mutant lines *pollb*, *reca1* and *why1why3*. (A) Plastid rearrangement breakpoint positions of the indicated *Arabidopsis* mutant lines. Heat maps depict each rearrangement as the intersection of the two genomic positions corresponding to the nucleotide on each side of the junction. Each tile represents a region

spanning 1 kb along each axis. Tile intensity represents the number of rearrangements per 10,000 plastid genomes. All rearrangements mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. The plastid large-single copy region (LSC), the first IR, and the small-single copy region (SSC) are depicted as a long grey bar, a red bar and a short grey bar, respectively. **(B)** Normalized amount of rearrangements per plastid genome displaying a microhomology (≥ 5 bp) (red) or not displaying a microhomology (blue) at their junction for each of the indicated mutant lines. **(C)** Proportion of total plastid rearrangements displaying a microhomology of given lengths, in bases, in each mutant line. Y axis represents the percentage of usage of each microhomology length. **(D)** Normalized amount of rearrangements corresponding to short-range inversions (< 50 bp, U-turn-like rearrangements, red) and short-range deletions/duplications (< 50 bp, blue) in each mutant line. Data were normalized on the total number of plastid genome reads. Del.: deletion, dup.: duplication. Two asterisks indicate a significant difference with p -value < 0.00001 using a chi-square test.

4.4.5 Whirly proteins, POLIB and RECA1 all act to maintain stability in the plastid genome

Genome maintenance is a tightly controlled process in which many proteins act in concert to repress the accumulation of DNA rearrangements. Consequently, the mutation of a single gene may not be sufficient to significantly destabilize the genome. We verified whether *POLIB*, *RECA1* and plastid Whirly genes have additive or synergistic effects by combining mutations for these genes in *Arabidopsis*. While the phenotype of a *reca1pollb* line is indistinguishable from that of WT plants (Fig. 39A), *why1why3pollb* is characterized by a severe growth retardation phenotype, in addition to yellow-variegation (Fig. 39A) [43,289]. *why1why3reca1* triple mutants show severe growth retardation as well and display white variegation and leaf distortion (Fig. 39A and Fig. 44A, page 145). Also, embryo lethality is observed in *why1why3reca1*, with 59% of seeds being unable to germinate (Fig. 44B, page 145). The quadruple mutant *why1why3pollbreca1* could not be isolated from the progeny of *why1why3^{-/-}pollb^{+/-}reca1^{+/-}* plants, nor from *why1why3^{-/-}pollb^{-/-}reca1^{+/-}* plants (Stouffer's Z test p -value < 0.05), suggesting that high levels of plastid genome instability leads to embryo lethality. While the overall pattern of ptDNA rearrangements in *reca1pollb* was similar to that of WT plants, an important increase was observed in both *why1why3pollb* and

why1why3reca1 (Fig. 39B and Supplemental Tables S20-22, <http://genome.cshlp.org/content/25/5/645/suppl/DC1>). Furthermore, a hotspot for rearrangements is present in the heat map of these two triple mutants, at the intersection of regions 86-101 kilobases (kb) and 68-100 kb. By contrast, the IR region from 101-108 kb, which encodes both plastid ribosomal RNAs, displays lower levels of instability.

Comparison of ptDNA rearrangements between *why1why3pollb* and *why1why3reca1* indicates that their amount of long-range rearrangements (≥ 1000 bp), arising with or without the use of microhomologies, are strikingly similar (Fig. 39D-E). This suggests that these types of rearrangements are unlikely to be responsible for the difference in phenotypes observed between the two lines. In contrast, rearrangements having occurred between regions separated by less than 1 kb were much more abundant in *why1why3reca1* than in *why1why3pollb* (Fig. 39C-D). *why1why3reca1* plastids also contain a 60-fold increase in U-turn-like rearrangements compared to WT plastids (Fig. 39F), which corresponds to approximately 60% of all short-range rearrangements present in this line. These results suggest that Whirly proteins and RECA1 both act to suppress the appearance of U-turn-like rearrangements in plastids. In addition, this high level of U-turn-like rearrangements could account for the severe phenotype observed in *why1why3reca1* plants (Fig. 39A).

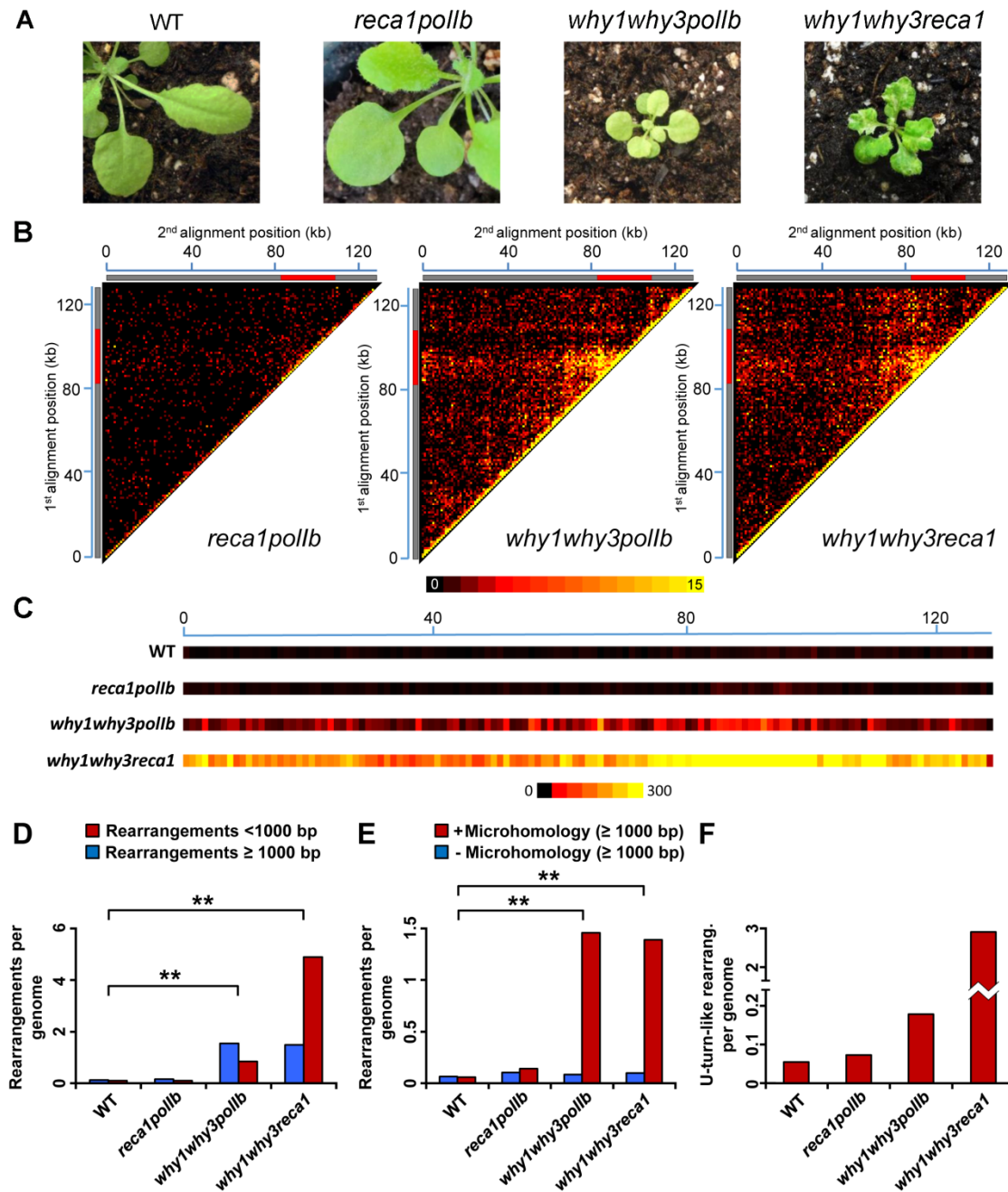


Figure 39. Global portrait of plastid genome rearrangements in *Arabidopsis* mutant lines *reca1pollb*, *why1why3pollb* and *why1why3reca1*. (A) Representative photographs of 21-d-old WT, *reca1pollb*, *why1why3pollb* and *why1why3reca1* *Arabidopsis* mutant plants. (B) Plastid rearrangement breakpoint positions of the indicated *Arabidopsis* mutant lines. Heat maps depict each rearrangement as the intersection of the two genomic positions corresponding to the nucleotide on each side of the junction. Each tile represents a region spanning 1 kb along each axis. Tile intensity represents the

number of rearrangements per 10,000 plastid genomes. All rearrangements mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. The plastid large-single copy region (LSC), the first IR, and the small-single copy region (SSC) are depicted as a long grey bar, a red bar and a short grey bar, respectively. **(C)** Short-range (<1000 bp) plastid rearrangement breakpoint positions of the indicated *Arabidopsis* mutant lines. Each heat map tile represents a region spanning 1 kb along the genome and the number of rearrangements per 10,000 plastid genomes. All rearrangements mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. **(D)** Normalized amount of short-range (<1000 bp) and long-range (≥ 1000 bp) rearrangements per plastid genome for each of the indicated mutant lines. **(E)** Normalized amount of long-range rearrangements (≥ 1000 bp) per plastid genome displaying a microhomology (≥ 5 bp) (red) or not displaying a microhomology (blue) at their junction for each of the indicated mutant lines. **(F)** Normalized amount of rearrangements corresponding to short-range inversions (<50 bp, U-turn-like rearrangements) in each mutant line. Data were normalized on the total number of plastid genome reads. Two asterisks indicate a significant difference with p -value < 0.00001 using a chi-square test.

4.4.6 U-turn-like rearrangements are associated to replication stress

The characteristics of the U-turn-like rearrangements identified above suggest that they might arise during replication, and that their formation could therefore be linked to replication stress. To test this hypothesis, we verified if a replication stress could be observed in *why1why3pollb* and *why1why3reca1*. Replication stress is defined here as the slowing of replication fork progression during DNA synthesis [152] which, in next-generation sequencing data, has been associated to a progressive, directional decrease in DNA copy number along the genome [303]. We therefore compared ptDNA coverage curves for each mutant line to the WT. Results indicate that the coverage in WT plants has a similar pattern to those previously reported in the literature for plastid genomes [304,305], and to that of all single and double mutant lines used in this study (Fig. 45, page 146). However, the pattern for ptDNA coverage appears quite different in *why1why3pollb* and *why1why3reca1* (Fig. 40A). Regression analysis within the large single-copy region (LSC, 0-84 kb) of *why1why3pollb* reveals a steeper slope than in the WT (Fig. 40B), suggesting that replication is affected in this mutant line and is unidirectional along the LSC, going from 84 kb toward the beginning of the genome. However, yet another pattern is observed in *why1why3reca1*, with two slopes

converging in the middle of the LSC in a manner consistent with bidirectional replication (Fig. 40B). Both of these patterns were also confirmed using quantitative PCR (Fig. 46, page 147). These results therefore suggest that the *pollb* and *reca1* mutations, when combined to the *why1why3* mutations, affect replication differently. Nevertheless, the steeper slopes in the LSC observed for *why1why3pollb* and *why1why3reca1* suggest a replication stress, which could be at the origin of the increase in short-range inversions observed in the plastid genome of both of these mutant lines (Fig. 39F). Interestingly, the essential role of RecA in replication fork reversal and restart was previously shown in bacteria and could account for the replication stress observed in *why1why3reca1*. This result thus supports a role for RECA1 in plastid DNA replication.

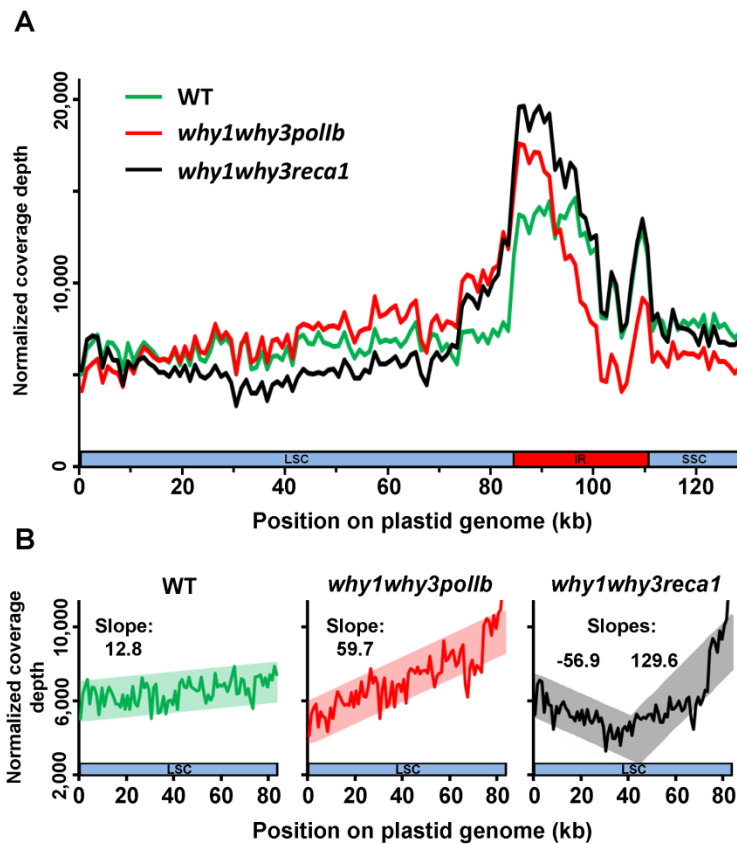


Figure 40. Plastid DNA sequencing coverage curves for *Arabidopsis* lines WT, *why1why3pollb* and *why1why3reca1*. (A) Plastid sequencing coverage of pools of 14-day-old *Arabidopsis* seedlings of the indicated genotypes. Positions were rounded down to 1 kb. All reads mapping to the plastid large inverted repeats

(IRs) were only assigned to the first IR. The plastid large-single copy region (LSC), the first IR, and the small-single copy region (SSC) are depicted as a long blue bar, a red bar and a short blue bar, respectively. Y axis represents the number of reads per 1,000,000 total plastid reads. **(B)** Regression analysis of the plastid large single-copy (LSC) sequencing coverage of WT, *why1why3pollb*, *why1why3reca1* seedlings. Positions were rounded down to 1 kb. Y axis represents the number of reads per 1,000,000 total plastid reads.

Since replication stress is generally associated to incomplete replication of chromosomes, pulse-field gel electrophoresis (PFGE) was used to visualize the distribution of plastid genomic molecules in the previous mutant lines. This confirmed an earlier study which demonstrated that, in contrast to the WT, no monomeric form of the chloroplast genome is observed in *reca1* mutant plants [210]. Similar to *reca1*, disappearance of the monomeric form is observed in *why1why3pollb* and *why1why3reca1* (Fig. 47, page 148). However, a smear of subgenomic molecules, likely associated with abortive rounds of replication, is instead observed in these mutants, supporting the hypothesis that the elevated levels of U-turn-like rearrangements in these lines could be the consequence of an ongoing replication stress. Note that subgenomic molecules are also observed in the *why1why3* mutant, suggesting that Whirly proteins are also important for replication.

4.5 Discussion

By sequencing DNA rearrangement junctions, we have been able to obtain a genome-wide portrait of DNA instability in the genomes of organelles and of a prokaryote. Unlike paired-end sequencing approaches used to detect rearrangements, our method is not affected by the distance over which the rearrangements occur. It therefore provides a view of rearrangements occurring both at short-range and long-range, while also allowing some characterization of the DNA repair mechanisms used (Fig. 48-49, pages 149-150). Indeed, it revealed that in *Arabidopsis* and human organelles, both microhomology-dependent and independent pathways constitute an important driving force of genome variation, with pathways using microhomologies being used slightly more often (Fig. 37B). Although illegitimate recombination has been known to occur in organelles [306,307], there was no

evidence that microhomology-independent repair, such as NHEJ, accounted for such a substantial proportion of organelle DNA rearrangements. Our approach also unveiled an unexpected pattern of genomic instability in organelles of both *Arabidopsis* and humans, with a strong propensity to generate rearrangements between closely spaced regions of the genome. Since DNA rearrangements are often associated to errors occurring during DNA repair [200,295,296], the high occurrence of short range rearrangements suggests that, in organelle genomes, inaccurate DNA repair takes place predominantly in the vicinity of the damaged DNA.

Our findings also reveal that, in organelles, a previously-unreported but frequent error-prone mechanism exists that most likely contributes to the restart of stalled replication forks. Indeed, we observed that U-turn-like rearrangements are particularly abundant among short-range DNA rearrangements in both *Arabidopsis* and human organelles. The finding that an apparent DNA replication stress correlates with the appearance of U-turn-like rearrangements in plastids suggests that they are initiated in a replication-dependent manner. Paused forks are unstable structures often formed in conditions of replication stress, which leads us to hypothesize that they act as a template for U-turn-like rearrangements. The tendency of these U-turn-like rearrangements to occur upstream and the presence of short inverted repeats at most of their junctions suggest that a small inverted repeat in the 3' end of the nascent strand would misanneal to its complement upstream on the opposite strand and lead to inaccurate fork restart (Fig. 41).

The fact that microhomologies favor U-turn-like rearrangements suggests that WHY1 and WHY3, which limit the accumulation of microhomology-dependent rearrangements, prevent this form of instability by repressing the misannealing of closely-spaced microhomologies. Also, the observation that single mutations of *RECA1* and *POLIB* have no visible impact on U-turn-like rearrangements but cause major changes only when combined to the mutations of *WHY1* and *WHY3* indicate the synergistic interaction between these genes and the Whirly genes. In the absence of Whirly proteins, replication stress could explain an increase in U-turn-like rearrangements, as observed in *why1why3pollb*. In contrast, we postulate that

replication stress alone does not likely account for the much larger increase observed in *why1why3recA1*. We propose instead that RECA1 is involved in a conservative repair pathway that directly competes with replication U-turns. Interestingly, the recombinase RecA has been shown to be essential in bacteria to efficiently bypass lesions and restart replication by promoting replication fork reversal [197,291,292]. It can thus be hypothesized that RECA1 also promotes lesion bypass and accurate fork restart in plastids, thereby limiting the accumulation of U-turn-like rearrangements. The severe phenotype observed in the *why1why3recA1* mutant also indicates that U-turn-like rearrangements have deleterious effects, and that their occurrence must be limited by multiple checkpoints.

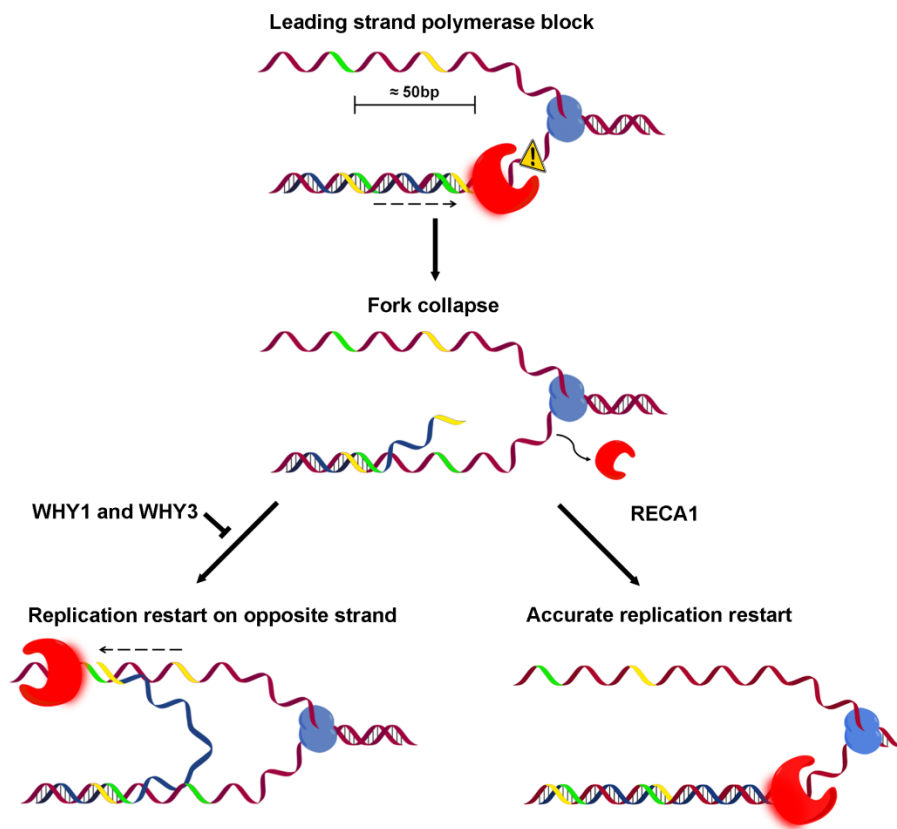


Figure 41. Model of microhomology-mediated U-turn-like inversions. The red shape represents the DNA polymerase and the blue shape represents the DNA helicase. Yellow and green strands represent inverted microhomologies. Impediments to replication fork progression might force the leading strand polymerase to pause and eventually unload. In the presence of RecA, the impediment may be bypassed to restart replication accurately. In the absence of

RecA and Whirly proteins, a microhomology located at the 3' end of the nascent strand can anneal to its complementary inverted repeat on the opposite strand and restart the replication fork on the wrong strand. The dotted arrow represents a multistep process.

Taken together, our results suggest a model for the generation of U-turn-like rearrangements in which progression of the leading strand polymerase is arrested during replication as a consequence of replication stress (Fig. 41). If the fork remains paused and eventually collapses, mechanisms such as RecA-dependent fork-reversal can accurately restart replication. In the absence of RECA1 and Whirly proteins, forks can restart inaccurately by the misannealing of a small inverted repeat in the 3' end of the nascent strand to its complement upstream on the opposite strand. Replication would then be re-initiated on the opposite strand until it either displaces the lagging strand or is ligated to its 5' end, resulting in the formation of a palindromic chromosome. Although this mechanism generates genomic rearrangements, it might serve as an alternative mechanism to restart replication forks. Whirly proteins would avert these rearrangements by binding to the single-stranded DNA and thus preventing misannealing of small inverted repeats.

Overall, our approach reveals that short-range rearrangements, and especially U-turn-like inversions, have been largely underestimated in organelle genomes. In this study, we report that high levels of U-turn-like rearrangements correlate with the appearance of the severe phenotype observed in *why1why3reca1* and to its 59% seed lethality. Since this type of genomic instability also constitute the majority of the DNA rearrangements occurring in human mitochondria, it warrants further investigation into its link to the development of some of poorly understood mitochondrial disorders. Mitochondrial genome instability has indeed been observed in many clinical disorders including Parkinson's disease [308,309], inclusion body myositis [310] and cancer [311]. In this regard, it will be interesting to evaluate if particular patterns of mitochondrial genomic instability are observed in the context of these disorders. Genomic rearrangements associated to inversions in the human nuclear genome have also been linked to leukemia, autism, and intellectual disability

[312,313]. It would therefore also be interesting to determine whether such rearrangements also occur through a U-turn-like mechanism and lead to the onset of such phenotypes.

4.6 Methods

4.6.1 Plant material and growth conditions

Arabidopsis thaliana (ecotype Columbia-0) mutant lines *reca1-2* (SALK_072979) [210], *pollb-1* (SALK_134274), *why1why3*, and *why1why3pollb-1* were reported previously [27,43]. The *reca1-1* (SALK_057982) mutant line was isolated from heterozygous seeds obtained from ABRC [314]. Seeds were sown on soil, vernalized for 3 days at 4°C, and grown under laboratory normal light (100 mmol m⁻² s⁻¹), at 22°C on a 16-h-day/8-h-dark cycle. Representative photographs were taken at 21 days. To determine statistical significance of *why1why3pollbreca1* lethality, chi-square tests were performed on progeny of *why1why3^{-/-}pollb^{+/-}reca1-1^{+/-}* and *why1why3^{-/-}pollb^{-/-}reca1-1^{+/-}* plants. Chi-square *p*-values were then combined using Stouffer's Z test [315].

4.6.2 DNA isolation and DNA-seq

For Col-0, *pollb*, *reca1*, *why1why3*, *reca1pollb*, *why1why3pollb* and *why1why3reca1*, total DNA was isolated from ≈400 mg pools of 14-day-old *Arabidopsis* plants using the cetyl trimethyl-ammonium DNA extraction protocol [249]. DNA was fragmented to ≈200-500 bp using S-Series Covaris according to Illumina's specifications. Libraries were prepared using the TruSeq DNA library preparation kit (Illumina) according to manufacturer's instructions. Efficient library generation was then assessed using a Bioanalyzer platform (Agilent) and Illumina MiSeq-QC run was performed. Sequencing was performed using an Illumina HiSeq 2000 using TruSeq SBS v3 chemistry at the Institute for Research in Immunology and Cancer's Genomics Platform (Université de Montréal). Cluster density was targeted at around 600 to 800 kilo clusters mm⁻². Sequencing data is available on the NCBI Sequence Read Archive (SRA) under the BioProject Number SRP051208 (Col-0: SRX883065, *pollb*: SRX813508, *reca1*:

SRX883066, *why1why3*: SRX883067, *reca1pollb*: SRX883068, *why1why3pollb*: SRX883069, *why1why3reca1*: SRX883070).

4.6.3 Publicly available Illumina whole-genome sequencing datasets

NCBI SRA accession numbers for paired-end Illumina whole-genome sequencing datasets for *Arabidopsis* ecotypes Ts-1 and Ws-2 are, respectively, SRX145018 and SRX145037 (submitted by the SALK Institute for Biological Studies). SRA accession numbers for paired-end Illumina whole-genome sequencing datasets for four human brain samples and four liver samples are, respectively, ERX385572, ERX385573, ERX385574, ERX385575 and ERX385576, ERX385577, ERX385578, ERX385579 (submitted by the Institute for Molecular Bioscience - The University of Queensland). SRA accession numbers for paired-end Illumina whole-genome sequencing datasets for *E. coli* are SRX154301, SRX154337, SRX154338 and SRX154342 (submitted by Indiana University).

4.6.4 Enrichment for reads with potential junctions

The Galaxy online software suite was used to develop a workflow that enriches reads spanning potential junctions from paired-end Illumina datasets (Fig. 48, page 149) [287]. As part of the workflow, quality filtering was performed to keep pairs for which both reads measure at least 40 bases, have an average quality of at least 20 and display no more than 50 bases outside of the quality range [316]. Using Burrows-Wheeler Aligner (BWA) [317], both reads from each pair were mapped individually against the whole reference genome (GRCh38 with alternate loci for *Homo sapiens* and Galaxy built-in reference genomes *Arabidopsis_thaliana_TAIR10* for *Arabidopsis thaliana* and *eschColi_K12* for *E. coli*) to remove pairs of read that do not contain a junction. This also allows the removal of nuclear plastid DNA sequences (NUPTs) and nuclear mitochondrial DNA sequences (NUMTs) that would otherwise be flagged as rearrangements. The alignment was performed without using a subsequence as a seed and parameters were set to a maximum edit distance of 7, a maximum number of gap opens of 1, a maximum number of gap extensions of 3, a mismatch penalty of 3, a gap open penalty of 11 and a gap extension penalty of 4. Pairs with at least

one unmapped read were then selected and the first 25 bases of both reads were mapped against the desired organelle genome (GenBank: AP000423.1 for *Arabidopsis* plastid, GenBank: Y08501.2 for *Arabidopsis* mitochondria, NCBI: NC_012920.1 for human mitochondria and Galaxy built-in reference genome eschColi_K12 for *E. coli*) using Bowtie for Illumina with default parameters [318]. Aligning the 25 first bases of both reads is critical, as rearrangement junctions would prevent the alignment of the whole rearranged read. The distance between each alignment position of a pair was then computed and pairs containing an unmapped read were discarded. Only pairs that fulfil one of the following conditions were then kept: 1- Each read of a pair is mapped in opposite orientation, 2- Each read of a pair is mapped in the same orientation with a distance of at least 3 bases. Finally, pairs were filtered to discard those that correspond to DNA fragments spanning junctions created by the different isoforms of the plastid genome or the circular nature of organelle DNA. Workflow statistics are listed in Tables III to VIII (pages 152-153). Enriched reads spanning potential junctions are provided in Supplemental Data File 1 (<http://genome.cshlp.org/content/25/5/645/suppl/DC1>) for all lines and organisms used in this study.

4.6.5 Analysis of rearranged reads

Sequences of reads containing a potential junction were aligned against the organelle genome (Positions 1 to 128214 of GenBank: AP000423.1 for *Arabidopsis* plastid, GenBank: Y08501.2 and JF729201.1 for *Arabidopsis* mitochondria, NCBI: NC_012920.1 for human mitochondria and NCBI: NC_000913.2 for *E. coli*) using BLAST+ [319] and the following parameters: “blastn -query potential_junctions.fasta -db organelle_genome.fasta -out output.txt -word_size 10 -max_target_seqs 2 -evaluate 0.0001 -outfmt 6”. Only reads with two alignments were kept for further analysis. Among the reads that output two alignments, the longest alignment was subtracted to the length of the read, and reads with values of 5 and less were discarded. BLAST+ outputs for rearrangements reads are provided in Supplemental Data File 2 (<http://genome.cshlp.org/content/25/5/645/suppl/DC1>) for all lines and organisms used in this study. To ensure that errors in wild-type plastid genome annotations do not generate false-positive rearrangements, the genomic sequence was confirmed using

BLAST+ for each locus at which rearrangements are more abundant than 5% of the average coverage. The reference locus sequence was used as the subject sequence to align the total reads of the samples. When rearrangements were found to be in homoplasmy, these rearrangements were considered as false-positives and were discarded. For *Arabidopsis* Col-0 mitochondria, two genome annotations exist (GenBank: Y08501.2 and JF729201.1). To remove false-positives, only reads considered as rearrangements on both annotations were kept. The following algorithms were then used to analyze the rearrangement junctions according to the alignment position output by BLAST+.

Distance measurement algorithm evaluates the difference between the reference genome positions of the 3' end of the first alignment and the 5' end of the second alignment. Overlap of alignments algorithm corresponds to the subtraction of the total read length to the sum of the lengths of both alignments. Rearrangements with an overlap of at least 5 bases are considered to have occurred through the use of microhomology, while the rest are assigned to the "No Microhomology" group. This high cut-off was used to ensure the stringency of microhomology calling. Inversion algorithm corresponds to rearrangements for which the reference genome positions from the 5' end to the 3' end are in ascending order for one alignment and in descending order for the other. All remaining rearrangements are treated as deletions or duplications. U-turn-like rearrangements are defined as those which present an inversion with a distance parameter of less than 50 bases. Downstream, same base and upstream U-turn-like rearrangements are determined by the position on the reference genome of the 5' end of the second alignment in relation to the 3' end and direction of the first alignment. The polarity of U-turn-like rearrangements identified in R2 reads has been inverted. Local mapping for read pairs associated to a single rearrangement were selected and aligned on the *Arabidopsis thaliana* (TAIR10) genome using the Golden Helix GenomeBrowse® visualization tool (Version 2.0.7). Representative results for a duplication, a deletion and an inversion are presented in Figure 50 (page 151).

To evaluate the proportion of U-turn-like rearrangements that occurred through the use of imperfect microhomologies, reads corresponding to this type of genomic rearrangement

were aligned against the plastid genome using BLAST+ and the following parameters: “blastn -query U-Turns.fasta -db plastid_genome.fasta -out output.txt -word_size 10 -max_target_seqs 2 -evalue 0.002 -outfmt 6 -penalty -1 -gap open 0 -gap extend 2”. The previous algorithms were then used to evaluate the proportion of rearrangements harboring a microhomology at the junction.

The combination of the previous workflow to enrich reads with potential junctions and these algorithms achieved 64.0% sensitivity for detection of reads containing junctions and 97.5% specificity for accurate rearrangement type calling. For sensitivity, 200 randomly chosen reads from the potential junctions obtained following the Galaxy workflow for the plastid DNA were blind-tested individually to assess if they correspond to a rearrangement or not. For specificity, 480 randomly chosen rearrangements analyzed by the algorithm were analyzed for accurate rearrangement type calling. For both sensitivity and specificity, the online BLAST interface (<http://blast.st-vi.ncbi.nlm.nih.gov/Blast.cgi>) was used with the lowest stringency settings.

4.6.6 Plastid sequencing coverage analysis

Pairs with both reads fully aligned against the reference genome using BWA during enrichment for potential junctions were filtered to keep only those mapping the plastid genome. Positions of each read were rounded down to the nearest kb and all reads mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. The number of reads mapping each 1 kb range was measured and normalized relative to 1,000,000 plastid reads.

4.6.7 Pulse-field gel electrophoresis (PFGE) analysis of the plastid genome

Chloroplasts were isolated from 21-day-old *Arabidopsis* plants as described previously [320]. Isolated chloroplasts were resuspended in homogenization buffer and then mixed 1:1 to 45°C 1% low melting point agarose in TE buffer and allowed to fix at 4°C. Isolated chloroplasts concentration were adjusted according to a low-cycle amplification of a DNA fragment of the hypothetical protein RF2 (YCF2) with the following primers: YCF2FOR, GAT CTC TGA GAG CTG

TTT CCG; YCF2REV, TGT TTC GCC TCT TAC TCG GAG. Agarose plugs were then soaked overnight at 50°C in lysis buffer (0.45 M EDTA pH 8.0, 1% (w/v) sarkosyl, 1 mg/mL proteinase K) and washed in storage solution (0.45 M EDTA pH 8.0, 1% (w/v) sarkosyl). Migration was performed in 1.5% agarose gel in 0.5X TBE buffer for 46 hours at 12°C using a Bio-Rad CHEF-DR® III system. Pulses switch times were set to 120 seconds at 5 V/cm using a 120 degrees angle. The gel was then soaked successively in 0.5X TBE buffer supplemented with 1 µg/mL RNase A for 2 hours at room temperature, in 0.5X TBE buffer supplemented with 5 µg/mL ethidium bromide for thirty minutes and finally in 0.5X TBE buffer to wash the gel. Southern hybridization was then performed as described previously [27] using a chloroplast probe amplified using the following primers: 49741FOR, CCT TAC GTA AAG GCC ACC CTA; 54551REV, TGG GAC GCA TAA CCG GAT ATG.

4.6.8 Quantitative PCR analysis of ptDNA levels

Total DNA was isolated from 14-day-old *Arabidopsis* Col-0, *why1why3pollb* and *why1why3reca1* plants using the cetyl trimethyl-ammonium DNA extraction protocol [249]. Primers used for qPCR were calibrated to ensure the amplification of a unique PCR product and efficiency between 1.90 and 2.05. Every reaction was carried out on biological and technical triplicates relative to the amplification of nuclear DNA. Primers sequences are as follows: 7660FOR, TGA TCC AGG ACG TAA TCC GGG AC; 7802REV, CGA ATC CCT CTC TTT CCC CTT CTC C; 45345FOR, TTG GCA ATT CCT CAG GGG CAG; 45525REV, TTG ACT ATT CCT CAA GCG CGC C; 81312FOR, AGC TAC CCA ATA CTC AGG GGA TCC; 81460REV, AAA TAG AAG CAG GGC GAC GCG; nucDNA-FOR, GTT GAA GCC TCC GTT CCC TGC TA; nucDNA-REV, CTC TTC CAC CGT GCA TGG CTT GT. The Power SYBR® Green PCR Master Mix (Applied Biosystems) was used according to manufacturer's instructions. qPCR experiments and analysis were carried out using LightCycler 480 (Roche) and the LightCycler 480 software version 1.5, respectively.

4.6.9 Quantitative PCR analysis of RECA1 expression in *reca1* mutants

Quantitative PCR analysis of *RECA1* expression was performed as described previously for 21-day-old plants [289]. Every reaction was carried out on biological and technical triplicates

relative to the amplification of beta tubulin. Primers used for qRT-PCR were calibrated to ensure the amplification of a unique PCR product and efficiency between 1.90 and 2.05. Primers sequences are as follows: RECA1FOR, GGT GGA GGC CTA CCA AAG GG; RECA1REV, GGT GGA GGC CTA CCA AAG GG; BetaTubFOR, TCG TTG GGA GGA GGC ACA GGT; BetaTubREV, GCT GAG TTT GAG GGT ACG GAA GCA G. The Power SYBR® Green PCR Master Mix (Applied Biosystems) was used according to manufacturer's instructions. qPCR experiments and analysis were carried out using LightCycler 480 (Roche) and the LightCycler 480 software version 1.5, respectively.

4.6.10 DATA ACCESS

All newly generated sequencing data is freely available on the NCBI Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under the BioProject Number SRP051208 (Col-0: SRX883065, *pollb*: SRX813508, *reca1*: SRX883066, *why1why3*: SRX883067, *reca1pollb*: SRX883068, *why1why3pollb*: SRX883069, *why1why3reca1*: SRX883070). The Galaxy workflow is freely available on Galaxy's Published Workflows section under the title "Rearrangement Junction Detection".

4.7 Acknowledgements

We thank M. Arteau, M. Valach and the Galaxy Team for assistance with the DNA-seq, PFGE and data processing, respectively. We also thank T. Vincent for his help in isolating mutant lines. This work was supported by scholarships from the Natural Sciences and Engineering Research Council of Canada (NSERC) to É.L., É.Z. and S.T.B., from the Fonds de Recherche du Québec – Nature et Technologies to S.T.B. and grants from NSERC to N.B.

4.8 Author contributions

É.Z. isolated plant DNA for sequencing and performed the PFGE and qPCR experiments. É.Z., É.L. and S.T.B. developed the bioinformatics approach. É.Z., É.L., S.T.B., S.T. and N.B. analyzed and interpreted the data. É.Z., É.L., S.T.B. and N.B. wrote the manuscript.

4.9 Supplementary data

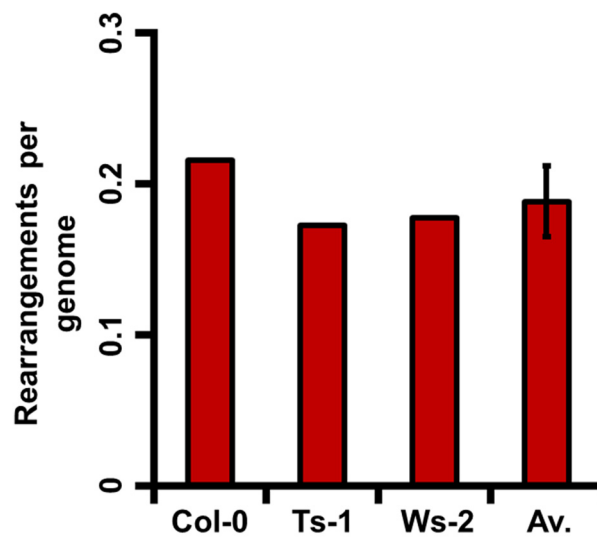


Figure 42. Level of plastid DNA rearrangements in three *Arabidopsis* ecotypes. Total number of rearrangements identified for each of the Col-0, Ts-1 and Ws-2 ecotypes. Y axis represents the number of rearrangements per 1,000,000 plastid reads. Av.: average of the three ecotypes. The error bar represents the standard deviation.

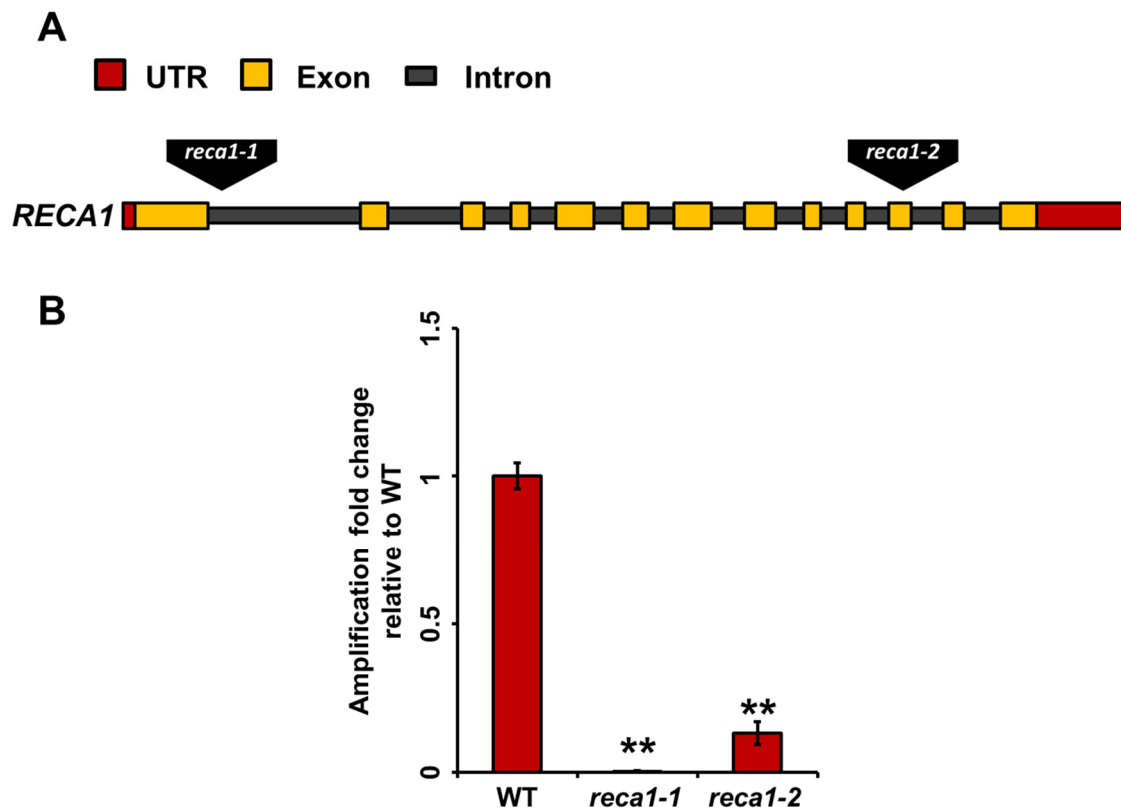


Figure 43. Characterization of the *reca1-1* and *reca1-2* T-DNA insertion mutant lines. (A) Schematic representation of the *RECA1* gene and insertion positions of the *reca1-1* and *reca1-2* insertions. (B) Quantitative PCR measurement of the *RECA1* expression levels for the *reca1-1* and *reca1-2* mutant lines relative to WT plants. Error bars represent the standard error of the mean of three biological replicates. Asterisks indicate a significant difference of a Student's *t* test *p*-value ≤ 0.01 with the WT.

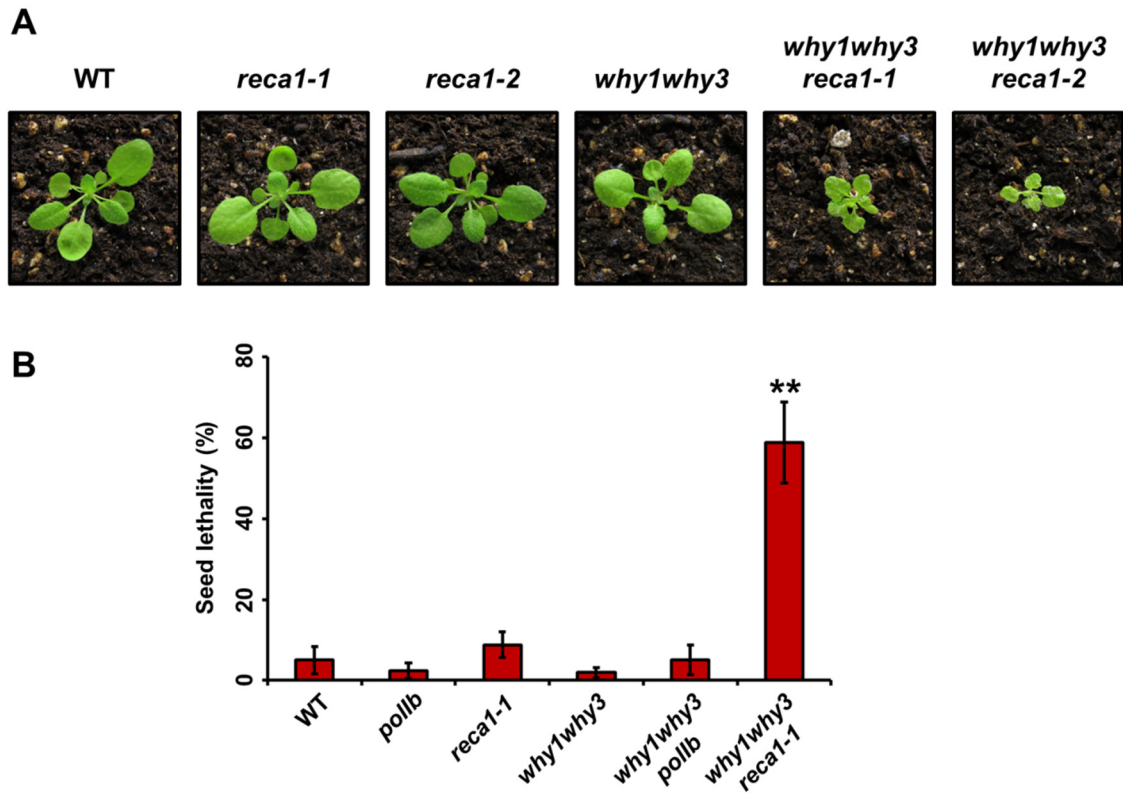


Figure 44. Characterization of the *why1why3reca1* phenotype. (A) Representative photographs of 21-day-old WT, *reca1-1*, *reca1-2*, *why1why3*, *why1why3reca1-1* and *why1why3reca1-2* *Arabidopsis* mutant plants. (B) Proportion of non-germinated seeds of WT, *pollb*, *reca1-1*, *why1why3*, *why1why3pollb* and *why1why3reca1-1*, four days after vernalization. Two asterisks indicate a significant difference of a Student's *t* test p -value ≤ 0.01 with the WT.

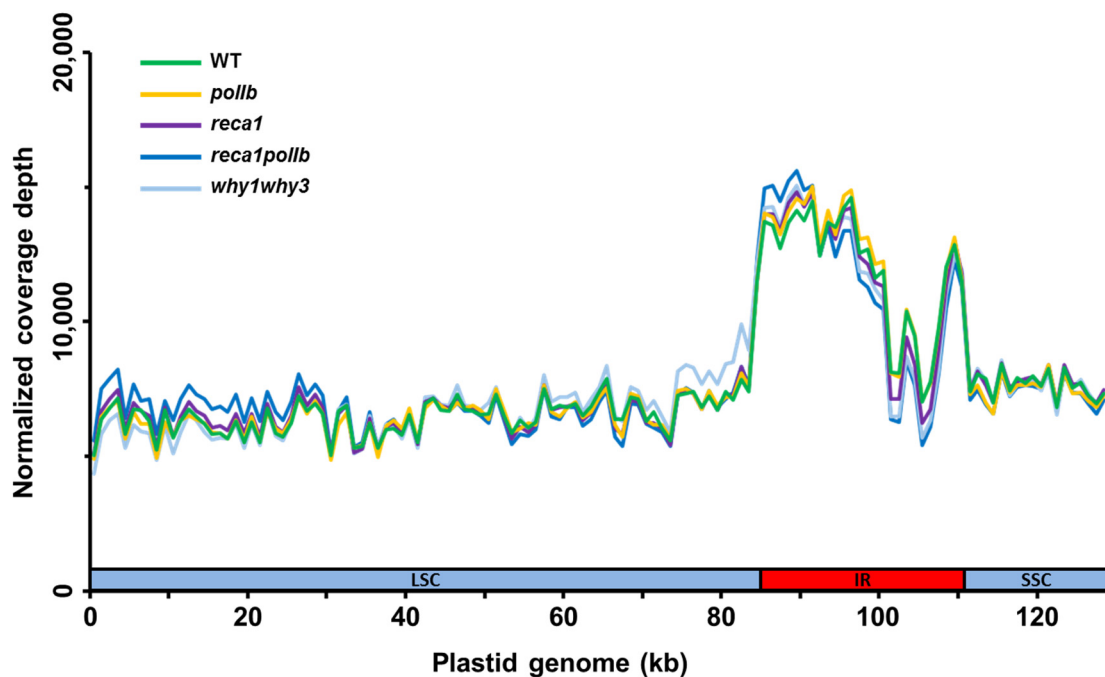


Figure 45. Plastid DNA sequencing coverage curves for WT, *pollb*, *reca1*, *reca1pollb* and *why1why3* *Arabidopsis* mutant plants. Plastid sequencing coverage of pools of 14-day-old *Arabidopsis* seedlings of the indicated genotypes. Positions were rounded down to 1 kb. All reads mapping to the plastid large inverted repeats (IRs) were only assigned to the first IR. Y axis represents the number of reads per 1,000,000 total plastid reads. The plastid large-single copy region (LSC), the first IR, and the small-single copy region (SSC) are depicted as a long blue bar, a red bar and a short blue bar, respectively.

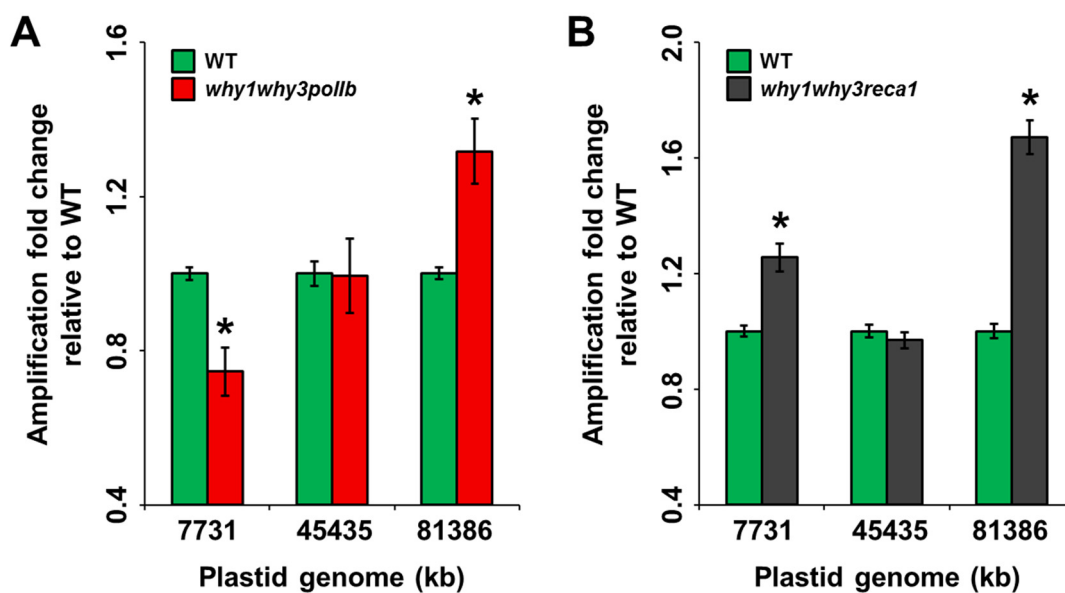


Figure 46. Plastid DNA quantification at three locations of the genome in WT, *why1why3pollb* and *why1why3reca1* plants. (A) Quantitative PCR measurement of ptDNA levels at three sites of the LSC in *why1why3pollb* relative to WT plants. **(B)** Quantitative PCR measurement of ptDNA levels at three sites of the LSC in *why1why3reca1* relative to WT plants. Error bars represent the standard error of the mean of three biological replicates. Asterisks indicate a significant difference of a Student's *t* test p -value ≤ 0.05 with the WT.

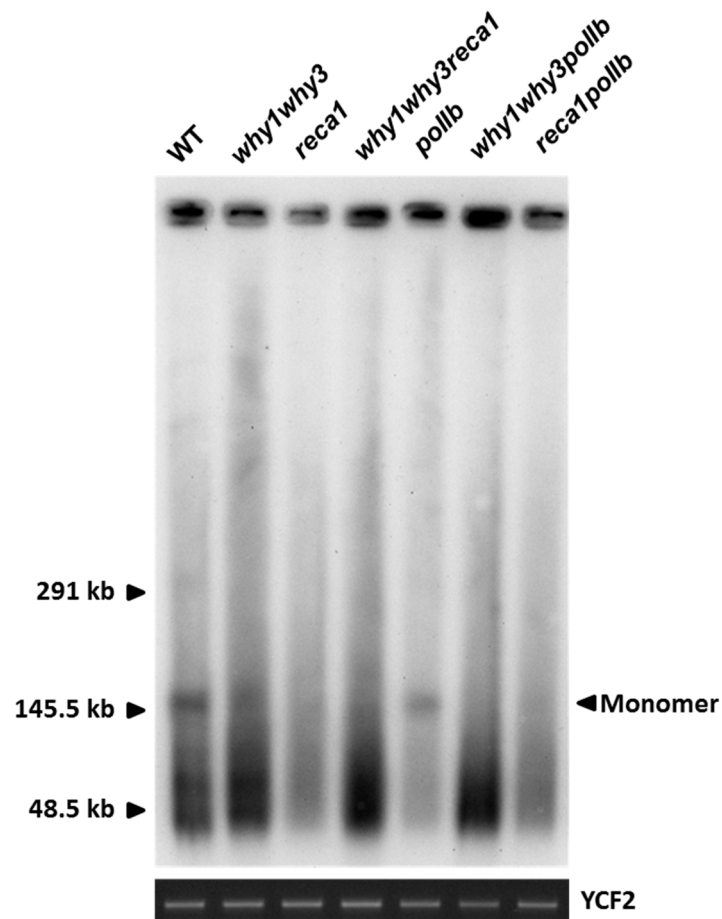


Figure 47. Visualization of the distribution of the various forms of the plastid genome. Pulsed-Field Gel Electrophoresis (PFGE) analysis of ptDNA in the indicated genotypes. The amount of DNA loaded in each well was equilibrated relative to the amplification of a *YCF2* fragment.

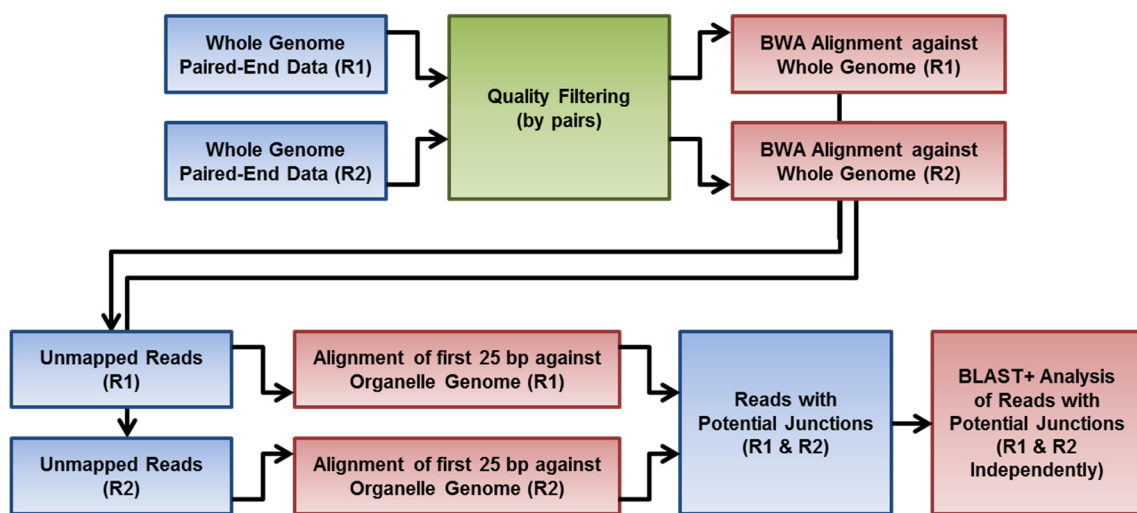


Figure 48. Schematic representation of the analysis workflow. Blue background boxes represent text manipulation steps, while the green and red backgrounds stand for quality filtering and mapping, respectively. Sequencing was performed on both ends of DNA fragments (R1 and R2). BWA: Burrows-Wheeler Aligner. R1 & R2: Paired-end sequencing read1 and read2.

	PCR and Southern Blot	Reporter Systems	Paired-End Analysis	Junction Analysis
Genome-wide	No <i>Primers and probe are specific to a single region</i>	No <i>Only the inserted region is observed</i>	Yes	Yes
Can be performed without genetic modification	Yes	No <i>Requires insertion of exogenous DNA into an organism's genome</i>	Yes	Yes
Allows comparison of samples	Yes	Yes	No <i>Detection is heavily dependent on DNA fragment length, which varies between samples</i>	Yes
Provides information about mechanism	Yes <i>Individual rearrangements need to be cloned and sequenced to obtain sequence</i>	Yes <i>Individual parameters can be varied to assess their importance</i>	No	Yes <i>Exact junction sequence provides some information about mechanism</i>
Provides a view of both short- and long-range rearrangements	No <i>Primers and probe are specific to a single region</i>	No <i>A single mechanism is observed at any time</i>	No <i>Bias toward long-range rearrangements</i>	Yes
Allows detection of short-ranged rearrangements	Yes	Yes	No <i>Detection of short-range rearrangements is limited by DNA fragment size</i>	Yes

Figure 49. Comparison of the techniques used to detect genome rearrangements. Details are provided when relevant.

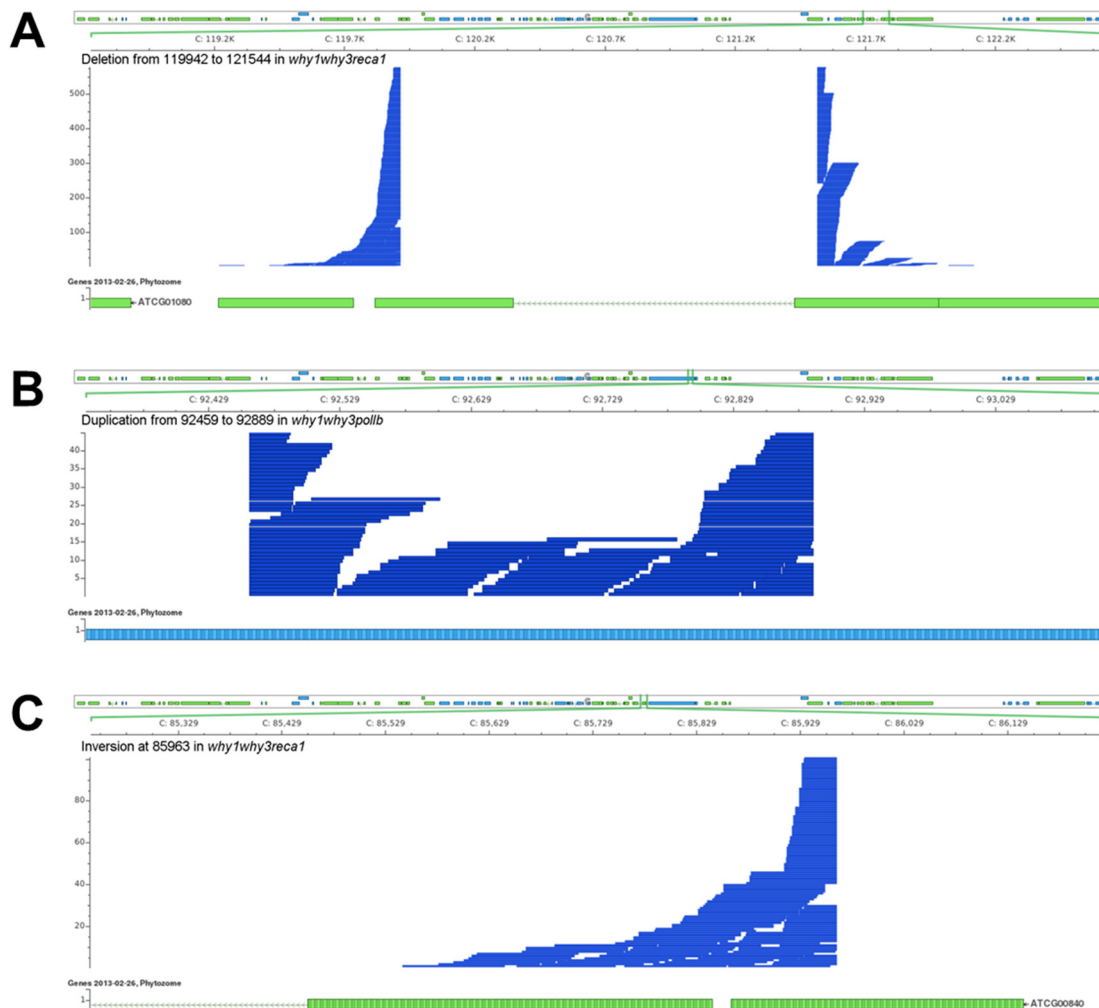


Figure 50. Local mapping of read pairs associated to specific genome rearrangements. The top box in each panel shows a map of the complete plastid genome on which genes on the forward strand are presented in blue and those on the reverse strand in green. The position encompassed by the zoom on the genome are presented below. The mapping of each read is presented as a blue rectangle. The bottom box shows a map of genes in the zoom. **(A)** Representative local mapping for a deletion. This deletion was observed 537 times in *why1why3reca1*. **(B)** Representative local mapping for a duplication. This duplication was observed 77 times in *why1why3pollb*. **(C)** Representative local mapping for an inversion. This inversion was observed 81 times in *why1why3reca1*.

Tableau III. Workflow Statistics for *Arabidopsis* plastid DNA rearrangements.

	Col-0	<i>pollb</i>	<i>reca1</i>	<i>reca1pollb</i>	<i>why1why3</i>	<i>why1why3</i> <i>pollb</i>	<i>why1why3</i> <i>reca1</i>
Read Pairs	31241719	32810928	30409865	31132957	31438219	32941816	31699029
Pairs with average quality >20	28771745	31343328	28967978	29738528	29944699	31308538	30156212
Plastid Pairs (%)	17,69	16,77	19,51	13,35	20,35	19,16	20,58

Tableau IV. Workflow Statistics for *Arabidopsis* ecotypes Ts-1 and Ws-2 plastid DNA rearrangements.

	Ts-1 (SRX145018)	Ws-2 (SRX145037)
Read Pairs	30012578	45521342
Pairs with average quality >20	26774195	40165795
Plastid Pairs (%)	9,89	9,54

Tableau V. Workflow Statistics for *Arabidopsis* mitochondria DNA rearrangements.

	Col-0
Read Pairs	31241719
Pairs with average quality >20	28771745
Mitochondria Pairs (%)	1,27

Tableau VI. Workflow Statistics for human brain mitochondria DNA rearrangements.

	erx385572	erx385573	erx385574	erx385575
Read Pairs	78908272	79738231	54905866	53255234
Pairs with average quality >20	54109701	54783744	52966448	51137017
Mitochondria Pairs (%)	0,55	0,55	0,61	0,64

Tableau VII. Workflow Statistics for human liver mitochondria DNA rearrangements.

	erx385576	erx385577	erx385578	erx385579
Read Pairs	74027979	75446145	47333976	46047382
Pairs with average quality >20	51040536	52250953	45679517	44520994
Mitochondria Pairs (%)	0,45	0,49	0,52	0,53

Tableau VIII. Workflow Statistics for *E. coli* DNA rearrangements.

	SRX154301	SRX154337	SRX154338	SRX154342
Read Pairs	2611112	2611112	2611112	2611112
Pairs with average quality >20	2611099	2611096	2611098	2611089
<i>E. coli</i> Pairs (%)	99,85	99,83	99,85	99,84

Chapitre 5. Discussion

5.1 L'instabilité génomique chloroplastique est associée à un déséquilibre photosynthétique

Afin de maintenir leur homéostasie rédox, il est essentiel pour les végétaux d'assurer un fonctionnement optimal de la PET. Les composantes de la PET étant encodées dans deux génomes distincts, une très fine régulation génique est nécessaire pour maintenir une stœchiométrie adéquate entre les différents constituants. L'instabilité génomique chloroplastique est donc susceptible de perturber ce fragile équilibre. Nous avons en effet démontré, par l'étude des mutants *why1why3pollb* et *why1why3reca1*, que de hauts niveaux de réarrangements génomiques dans l'ADNpt corrèlent avec une diminution de la capacité photosynthétique (Fig. 14 et 34, pages 59 et 108). Une corrélation similaire est également observée chez *Arabidopsis* lors d'un traitement à la CIP, un agent induisant des DSBs et des réarrangements génomiques dans l'ADN des organites (Fig. 26, page 88).

L'instabilité génomique chloroplastique est susceptible de déséquilibrer la PET de nombreuses façons. Tout d'abord, plusieurs des réarrangements génomiques observés chez *why1why3pollb* et *why1why3reca1* sont générés par un mécanisme dépendant de microhomologies associé, vraisemblablement, au processus de MMBIR. Puisque ce mécanisme génère une multitude de réarrangements génomiques complexes, dont des duplications et des délétions [200], il est possible que le nombre de copies de plusieurs gènes chloroplastiques soit modifié et que la stœchiométrie entre les composantes de la PET s'en retrouve bouleversée. Par ailleurs, les mutants *why1why3pollb* et *why1why3reca1* démontrent un probable stress réplcatif chloroplastique (Fig. 40, page 131). Des niveaux d'ADN inégaux entre différentes régions du génome perturbent presque assurément la stœchiométrie entre les composantes de la PET. Finalement, puisque les réarrangements génomiques altèrent la séquence nucléotidique, il est possible que ceux-ci inactivent des gènes, les tronquent ou génèrent des cadres ouverts de lecture chimériques. Des formes alternatives de nombreux transcrits chloroplastiques sont d'ailleurs observées chez *why1why3*, mais pas chez les plantes de type sauvage (A. Maréchal, résultats non-publiés). De tels transcrits aberrants risquent donc d'entraver le bon fonctionnement de la PET.

L'étude du mutant *why1why3pollb* et de plantes sauvages traitées à la CIP a aussi révélé que le déséquilibre photosynthétique, associé aux réarrangements génomiques chloroplastiques, semble être à l'origine d'un grave stress oxydatif dans le chloroplaste (Fig. 15 et 16, pages 61 et 63). Plusieurs types de ROS distincts peuvent être générés par la PET. Alors que le PSI est principalement associé à la production d'anions superoxydes et de peroxyde d'hydrogène, le PSII et ses complexes d'antennes sont essentiellement liés à la production d'oxygène singulet [revu dans 80]. Parmi les types de ROS associés aux réarrangements génomiques chloroplastiques, le NBT nous révèle la présence d'anions superoxydes [248]. Quant au H₂DCFDA, celui-ci nous informe peu quant aux types de ROS chloroplastiques présents, cette molécule étant peu spécifique [321-323]. Toutefois, puisque les réarrangements génomiques affectent probablement plusieurs composantes de la PET, plusieurs types de ROS contribuent vraisemblablement au déséquilibre rédox observé.

Chez *why1why3pollb* et *why1why3reca1*, ce déséquilibre rédox est associé à une complexe reprogrammation génique nucléaire qui promeut une résistance aux stress oxydatifs subséquents (Fig. 20, 22 et 34, pages 70, 75 et 108). Ceci implique que des voies de signalisation rétrograde sont actives entre le chloroplaste et le noyau de ces deux lignées. Il est toutefois difficile d'identifier lesquelles des voies de signalisation rétrograde participent à la reprogrammation génique de ces deux mutants. Néanmoins, il est vraisemblable que plus d'une voie de signalisation rétrograde soient activées chez *why1why3pollb* et *why1why3reca1*. Par exemple, il est probable que les réarrangements génomiques chloroplastiques altèrent l'état rédox de la PET, celui-ci étant à l'origine d'une voie de signalisation rétrograde [revu dans 87]. Également, il est possible que les voies de signalisation rétrograde dépendantes du PAP, du MEcPP et du β -cyclocitral participent à la reprogrammation génique des mutants *why1why3pollb* et *why1why3reca1*. Les patrons transcriptomiques de lignées d'*Arabidopsis* accumulant de hauts niveaux de PAP et de β -cyclocitral ne présentent toutefois qu'une faible similarité avec le patron de *why1why3pollb* [103,104] (E. Zampini, résultat préliminaire non-publié). Quant au MEcPP, aucune analyse de son transcriptome n'est disponible pour le moment. La participation de ces métabolites à la

reprogrammation génique associée aux réarrangements génomiques chloroplastiques demeure donc inconnue pour l'instant.

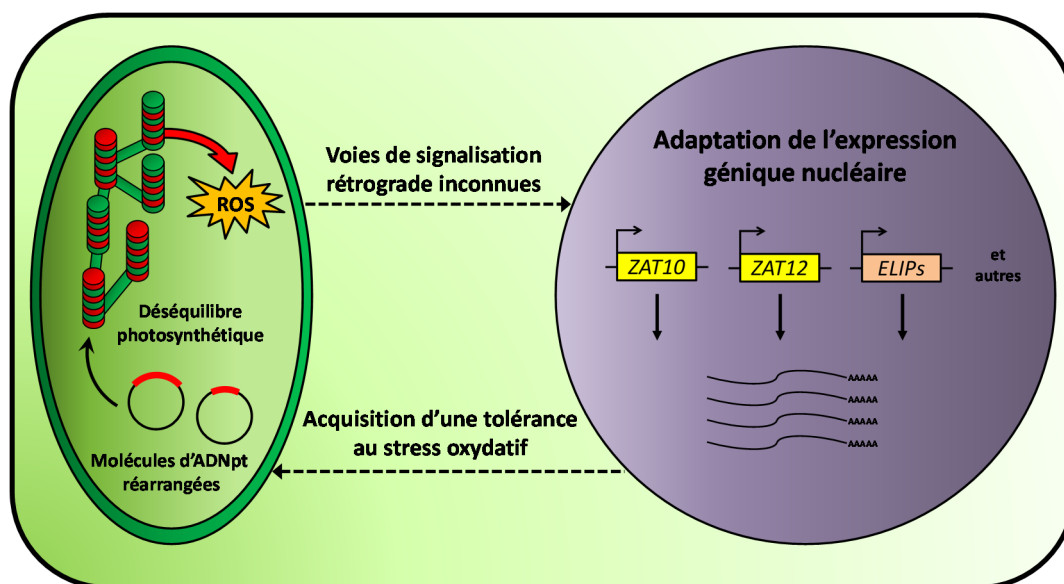


Figure 51. L'instabilité génomique chloroplastique est associée à un déséquilibre rédox et à une reprogrammation génique nucléaire. L'instabilité génomique chloroplastique provoque un déséquilibre de la PET qui, à son tour, est à l'origine d'une grande production de ROS. Des voies de signalisation rétrograde inconnues reprogramment ensuite le patron d'expression génique nucléaire et permettent une adaptation au stress oxydatif.

D'autre part, le mutant *why1why3pollb* présente un patron de reprogrammation génique hautement similaire à plusieurs conditions connues de stress oxydatif (Fig. 20, page 70). Ceci suggère que des voies de signalisation rétrograde dépendantes des ROS sont activées en conséquence des réarrangements génomiques chloroplastiques. Cette hypothèse est d'ailleurs supportée par l'absence de reprogrammation génétique dans des conditions de faible intensité lumineuse chez *why1why3pollb* (Fig. 21, page 73). Plusieurs évidences suggèrent que les protéines EXECUTER1 et 2 sont des acteurs clés dans l'acheminement au noyau de la signalisation dépendante de l'oxygène singulet [97,98]. De manière intéressante, le mutant *ex2* (*executer2*) présente une hypersensibilité à la CIP, suggérant qu'EXECUTER2 participe à l'adaptation des plantes à l'instabilité génomique chloroplastique (E. Lepage, E.

Zampini, résultat préliminaire non-publié). La génération du mutant *why1why3pollbex2* permettrait possiblement de répondre à cette question.

Globalement, cette étude a permis d'établir que l'instabilité génomique chloroplastique est associée à une déstabilisation de la PET et à un déséquilibre rédox chloroplastique (Fig. 51). Il a également été démontré que ce déséquilibre rédox est à l'origine de voies de signalisation rétrograde reprogrammant le patron d'expression génique nucléaire et assurant l'acquisition d'une tolérance aux stress oxydatifs subséquents.

5.2 Les protéines WHY1, WHY3, POLIB et RECA1 travaillent toutes au maintien de la stabilité génomique chloroplastique

Les patrons d'instabilité génomique obtenus au chapitre quatre suggèrent un important rôle des protéines Whirly, de l'ADN polymérase POLIB et de la recombinaise RECA1 pour le maintien de la stabilité génomique chloroplastique. Les réarrangements génomiques analysés dans cette étude ont été classés selon la distance séparant les deux côtés de la jonction séquencée, une distance de plus de 1000 nucléotides définissant le groupe des réarrangements formés à grande distance. Une distance de 1000 nucléotides permet vraisemblablement de discriminer les réarrangements génomiques associés à la réparation d'un DSB de ceux formés par le processus de réplication de façon indépendante d'un DSB. Bien que la réparation des DSBs contribue certainement aux réarrangements formés à courte distance, il semble probable que cet apport soit modeste, contrairement à la très grande proportion de réarrangements génomiques de type U-turn (voir aussi section 5.3.2, page 164). Quant aux réarrangements à grande distance (>1000 bp), ceux-ci sont probablement presque exclusivement associés à la réparation des DSBs, les réarrangements de type U-turn se formant principalement à très courte distance (Fig. 37, page 123). Les patrons d'instabilité génomique à grande distance (>1000 bp) nous informent donc probablement sur les mécanismes de réparation des DSBs.

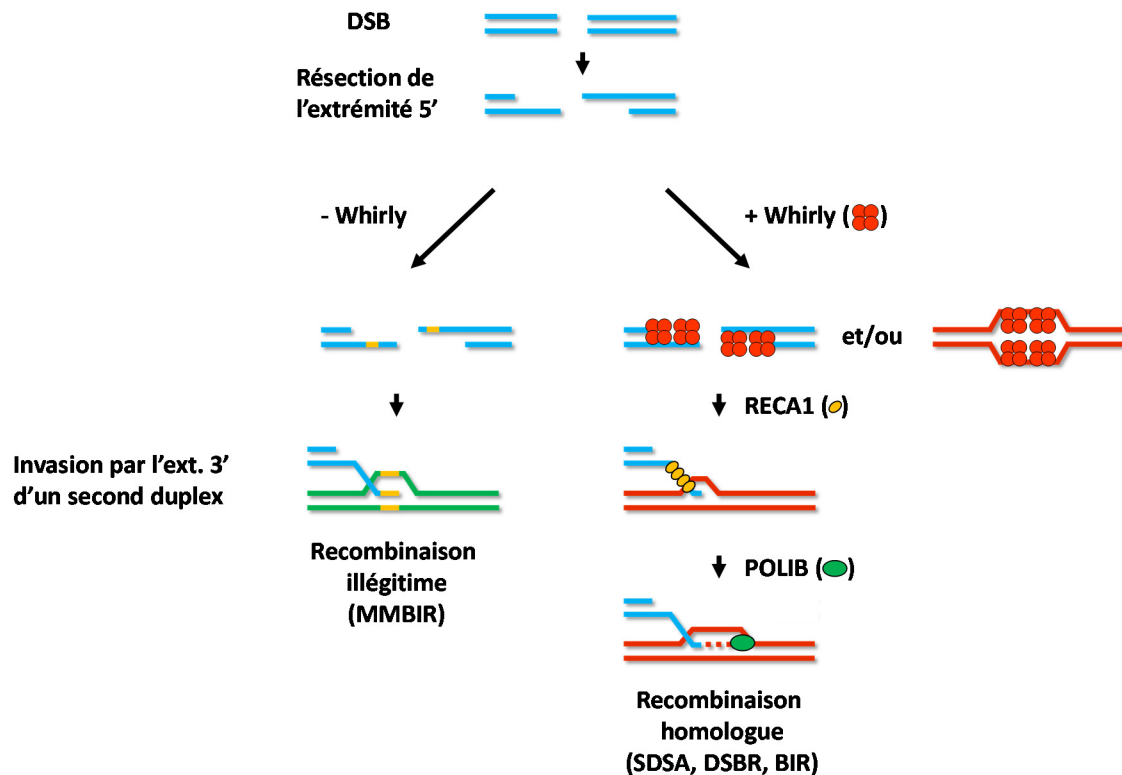


Figure 52. Les protéines Whirly, POLIB et RECA1 favorisent la stabilité génomique chloroplastique. Lorsqu'un DSB est généré, la liaison des protéines Whirly à l'ADNss limiterait les appariements non-spécifiques de manière à favoriser les voies de réparation de l'ADN appropriées. Lors de la réparation des DSBs par recombinaison homologue, la protéine RECA1 assurerait l'échange de brins avec un duplex d'ADN homologue (duplex rouge) et l'ADN polymérase POLIB synthétiserait la séquence d'ADN perdue lors du bris. Le duplex d'ADN vert représente un duplex d'ADN non-apparenté. Les tirets jaunes représentent des séquences de microhomologie.

Ces patrons de réarrangements à grande distance ont révélé une accumulation de hauts niveaux de réarrangements génomiques dépendants de microhomologies chez *why1why3*, *why1why3pollb* et *why1why3reca1*. La similitude des patrons des mutants *why1why3pollb* et *why1why3reca1* est d'ailleurs frappante (Fig. 38 et 39, pages 126 et 129). Ceci suggère que les mutations *pollb* et *reca1* affectent de manière similaire la réparation des DSBs lorsque combinées aux mutations *why1why3*. Tel que discuté au chapitre un, les mécanismes de MMBIR et de recombinaison homologue sont probablement en compétition pour la réparation des DSBs (Fig. 12 et 13, pages 41 et 46). L'accumulation très importante de

réarrangements génomiques dépendants de microhomologies chez *why1why3pollb* et *why1why3reca1* suggère donc qu'un mauvais fonctionnement de la recombinaison homologue aggrave fortement l'instabilité génomique lorsque les voies de réparation dépendantes de microhomologies ne sont plus inhibées. Ceci suggère un modèle où les protéines POLIB et RECA1 travaillent de concert avec les protéines Whirly à la réparation adéquate des DSBs. Alors que POLIB et RECA1 favorisent probablement la réparation des DSBs par recombinaison homologue, les protéines Whirly, quant à elles, semblent réprimer les mécanismes de réparation sujets aux erreurs dépendants de microhomologies (Fig. 52). Il est d'ailleurs intéressant de noter que la mutation des protéines Whirly chloroplastiques semble être un prérequis à l'accumulation de hauts niveaux d'instabilité génomique. En effet, aucune augmentation notable de l'instabilité génomique n'est observée chez les mutants *reca1*, *pollb* et *reca1pollb* comparativement aux plantes de type sauvage (Fig. 38 et 39, pages 126 et 129).

Il est toutefois impossible de confirmer pour l'instant que les protéines RECA1 et POLIB participent toutes les deux à la même voie de réparation des DSBs, par recombinaison homologue. En effet, la caractérisation du double mutant *reca1pollb* n'a révélé aucune interaction génétique entre ces deux gènes, hormis une forte hypersensibilité à la CIP (Fig. 32, page 103). Il serait normalement attendu que la participation de *RECA1* et de *POLIB* à une même voie de réparation résulte en une relation d'épistasie entre les deux gènes, c'est-à-dire que les conséquences de la double mutation soient identiques à chacune des simples mutations. Ce type de relation génétique survient lorsque chacune des simples mutations est suffisante à abolir une voie biologique. Il semble néanmoins possible que *RECA1* et *POLIB* travaillent au sein d'une même voie de recombinaison homologue, les mutations de chacun de ces gènes n'étant possiblement pas suffisantes pour abolir entièrement l'activité de recombinaison homologue chloroplastique. En effet, les gènes *RECA2* et *POLIA* pourraient partiellement compenser la perte d'activité de *RECA1* et *POLIB*, respectivement. De plus, la très grande similitude des patrons d'instabilité génomique des triples mutants

why1why3pollb et *why1why3reca1* suggère que les mutations *reca1* et *pollb* affectent de manière similaire la réparation des DSBs.

D'autre part, il est possible que de hauts niveaux d'instabilité génomique résultent plutôt d'une augmentation de la fréquence de formation de DSBs chez les lignées mutantes étudiées. Ceci semble néanmoins peu probable, les niveaux de réarrangements indépendants de microhomologies n'étant que peu affectés chez ces lignées. En effet, il serait attendu qu'une augmentation de la fréquence de DSBs résulte en une utilisation accrue de tous les mécanismes de réparation sujets aux erreurs et non uniquement de ceux dépendants de microhomologies.

5.3 Le processus de réplication menace l'intégrité génomique des chloroplastes et des mitochondries

5.3.1 : Les mécanismes de redémarrage des fourches de réplication bloquées

Afin de pouvoir discuter d'un mécanisme à l'origine des réarrangements de type U-turn et du rôle que pourraient jouer les protéines RECA1, WHY1 et WHY3 dans la prévention de ce type d'instabilité génomique, il est essentiel d'introduire auparavant les causes prohibant la progression des fourches de réplication ainsi que les mécanismes assurant leur redémarrage.

Lors de leur progression, les fourches de réplication rencontrent une multitude d'obstacles pouvant ralentir et bloquer leur progression. Les difficultés rencontrées pendant le processus de réplication sont collectivement regroupées sous le nom de stress réplcatif. Ceci inclut de nombreuses lésions bloquant les ADN polymérases, telles que des bases azotées oxydées, des dimères de nucléotides induits par les rayons UV-B et des ribonucléotides insérés dans la séquence d'ADN [revu dans 152]. Les structures secondaires de l'ADN, les ARNs polymérases en cours de transcription et la déplétion du pool de désoxynucléotides constituent également des obstacles pouvant stopper la progression des ADN polymérases. Puisque les fourches de réplication immobilisées sont des structures hautement instables, celles-ci doivent absolument être remises en marche afin de préserver l'intégrité génomique et de

compléter la ronde de réplication avortée. Les obstacles cités ci-haut n'ont toutefois pas les mêmes conséquences, dépendamment du brin sur lequel ils sont rencontrés. En effet, il est généralement accepté que l'inhibition de la progression des ADN polymérases constitue un défi bien moins important sur le brin discontinu, la synthèse d'ADN pouvant être ré-initiée en aval de l'obstacle sur ce brin [revu dans 225,324]. La situation est toutefois bien différente sur le brin continu, où de complexes mécanismes sont nécessaires au redémarrage de la réplication.

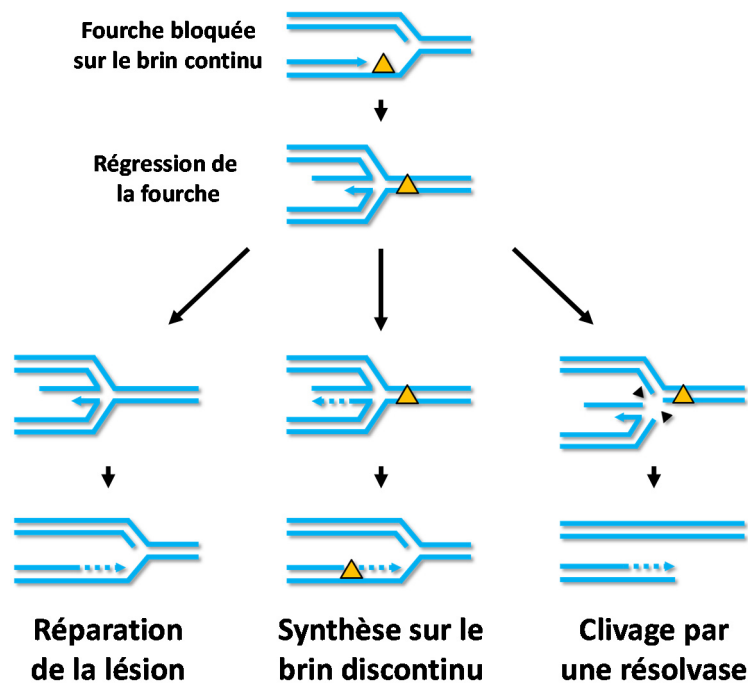


Figure 53. La régression des fourches de réplication bloquées et les mécanismes de redémarrage de la réplication. Lors de la rencontre d'une lésion bloquant la progression de l'ADN polymérase sur le brin continu, la régression de la fourche de réplication peut être catalysée de manière à former une structure en « chickenfoot ». Trois principaux mécanismes assurent ensuite le redémarrage de la réplication : 1) le repositionnement de la lésion dans un duplex d'ADN peut faciliter la réparation, 2) la synthèse du brin continu peut se poursuivre temporairement sur le brin discontinu et 3) la structure en « chickenfoot » peut être clivée par une résolvasse et la réplication redémarrée par BIR (non illustré sur la figure). Les pointillés représentent de l'ADN synthétisé suite à l'interruption de la ronde de réplication.

Deux principaux mécanismes semblent assurer le redémarrage des fourches de réplication bloquées sur le brin continu chez *E. coli*. Il a longtemps été considéré que seul le brin discontinu avait la capacité d'initier la polymérisation d'ADN ailleurs qu'à l'endroit de l'origine de réplication. De récentes évidences suggèrent toutefois que le brin continu possède également cette capacité et que, sur ce brin, la synthèse d'ADN peut être ré-initiée de façon discontinue en aval d'une lésion bloquant l'ADN polymérase [325]. La région d'ADNss générée par cette synthèse discontinue pourrait ensuite être comblée par recombinaison.

Le second mécanisme assurant le redémarrage des fourches de réplication bloquées sur le brin continu est toutefois beaucoup mieux établi. Celui-ci est qualifié de régression de la fourche (*fork regression*, également appelé déroulement de la fourche) et implique le ré-appariement des deux brins matrices simultanément à l'appariement des deux brins nouvellement synthétisés. Cette régression de la fourche de réplication génère une structure d'ADN très similaire à une jonction de Holliday, également appelée « chickenfoot », qui constitue le substrat de nombreuses voies de redémarrage de la réplication [revu dans 324] (Fig. 53). Parmi celles-ci, le repositionnement de la lésion au sein d'un duplex d'ADN pourrait faciliter les processus de réparation. Également, puisqu'une lésion sur le brin continu n'inhibe pas la synthèse d'ADN sur le brin discontinu, il est probable que l'ajout de nucléotides se poursuive sur une certaine distance en aval de la lésion sur le brin discontinu. Lors de la formation de la structure en « chickenfoot », l'appariement des deux brins nouvellement synthétisés permet ainsi l'élongation du brin continu en se servant du brin discontinu comme matrice. Cette élongation du brin continu permet, suite au ré-enroulement de la fourche, la progression de la synthèse d'ADN au-delà de la lésion. Finalement, les structures en « chickenfoot » peuvent également être clivées par une résolvasse, générant un DSB pouvant être réparé par BIR dans une ultime tentative de redémarrer la réplication lorsque les autres mécanismes ont échoués. Bien que la machinerie assurant la régression des fourches de réplication diffère, des mécanismes semblables

assurent néanmoins le redémarrage des fourches de réplication bloquées chez les eucaryotes [326].

5.3.2 Évidences d'un mécanisme de redémarrage des fourches de réplication de type U-turn dans le chloroplaste et dans la mitochondrie

Les réarrangements génomiques qualifiés de U-turn au chapitre quatre semblent représenter une très importante source d'instabilité génomique, tant dans les mitochondries humaines que dans les chloroplastes et les mitochondries d'*Arabidopsis* (Fig. 37, page 123). Dans ce chapitre, il est suggéré que suite à l'effondrement d'une fourche de réplication, l'appariement entre une microhomologie située à l'extrémité 3' du brin continu et une séquence complémentaire sur le brin opposé peut mener à la ré-initiation de la synthèse d'ADN sur le mauvais brin matrice (Fig. 41, page 134). Il résulterait de cette ré-initiation erronée de la réplication une inversion à très courte distance. Les jonctions analysées par séquençage de nouvelle génération dans cette étude ne permettent toutefois pas d'affirmer qu'un tel mécanisme alternatif de redémarrage des fourches de réplication est à l'origine de ce type d'instabilité génomique. Néanmoins, plusieurs de nos propres observations renforcent le modèle proposé. De plus, les réarrangements génomiques qualifiés de U-turn dans notre étude partagent de nombreuses similitudes avec d'autres mécanismes d'instabilité génomique documentés dans la littérature scientifique. Il sera donc ici question des évidences renforçant le modèle proposé au chapitre précédent.

Tout d'abord, il convient de mentionner que les inversions se produisant à très courte distance (<50 bp) sont largement surreprésentées comparativement aux délétions et aux duplications. En effet, plus de 96 % des réarrangements à très courte distance (<50 bp) sont des inversions dans le chloroplaste des plantes de type sauvage. Néanmoins, les très courtes inversions (<50 bp) observées dans notre étude pourraient évidemment constituer le produit de mécanismes de réparation des DSBs dépendants de très courtes microhomologies. Le processus de MMBIR génère cependant une multitude de produits réarrangés distincts, dont des duplications, des délétions et des inversions [200]. Quant au MMEJ, celui-ci est essentiellement associé à des délétions [188]. Il est donc peu plausible que de tels

mécanismes ne génèrent qu'une quasi-totalité d'inversions. D'autre part, alors que les inversions sont générées par le biais de séquences répétées et inversées, les duplications et les délétions sont, quant à elles, associées à des répétitions ayant la même orientation. Par conséquent, il est possible que la quasi-totalité des très courtes inversions (<50 bp) provienne simplement d'une surreprésentation de microhomologies d'orientation inversée dans le génome. Toutefois, bien que les microhomologies inversées (<50 bp) sont plus abondantes que celles de même orientation dans le génome chloroplastique (Fig. 54), la différence de proportion n'apparaît pas suffisante pour justifier la surreprésentation de très courtes inversions (<50 bp). Ceci suggère qu'un mécanisme particulier est probablement à l'origine des réarrangements de type U-turn.

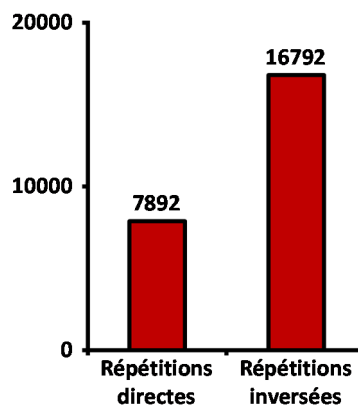


Figure 54. Distribution des microhomologies peu espacées dans le génome chloroplastique d'*Arabidopsis*. Comparaison entre les microhomologies (≥ 5 bp) de même orientation ou d'orientation inverse et séparées par moins de 50 bp.

Par ailleurs, le mécanisme que nous suggérons pour la formation des réarrangements de type U-turn nécessite un arrêt de la progression de l'ADN polymérase sur le brin continu. Une accumulation accrue de ce type d'instabilité génomique peut donc être logiquement associée à : 1) une augmentation de la fréquence des fourches de réplication bloquées et 2) une incapacité à redémarrer efficacement la réplication lors de tels blocages. Il semble d'ailleurs probable que de telles éventualités soient rencontrées chez les mutants *why1why3pollb* et *why1why3reca1*, ces deux lignées présentant, respectivement, des niveaux de

réarrangements de type U-turn chloroplastiques d'environ trois et 60 fois ceux rencontrés dans une plante de type sauvage. En effet, puisqu'un stress réplcatif semble être observé chez ces deux lignées (Fig. 40, page 131), la fréquence de fourches de réplication bloquées est vraisemblablement augmentée dans leurs chloroplastes. Par ailleurs, la très forte accumulation de réarrangements de type U-turn chez *why1why3recA1* suggère en plus un rôle pour ces protéines dans la stabilisation et dans le redémarrage des fourches de réplication. Ceci semble d'ailleurs en accord avec les rôles importants de la recombinaise bactérienne RecA et des protéines de liaison à l'ADNss dans la réplication de l'ADN bactérien (voir section 5.3.3).

Finalement, des inversions génomiques similaires sont également recensées chez d'autres organismes et confortent le modèle suggéré ici. Chez *E. coli*, deux études ont démontré que des répétitions inversées de quelques centaines de nucléotides peuvent mener à la formation d'inversions génomiques dans un processus vraisemblablement associé à la réplication et, tout comme dans notre étude, indépendant de RecA [327,328]. De manière intéressante, il avait été noté qu'une diminution de la distance entre les deux répétitions inversées résulte en une augmentation exponentielle du taux de « fusion » des deux séquences répétées et inversées [328]. Quoique dans cette étude la distance entre les répétitions soit beaucoup plus grande que pour les réarrangements de type U-turn observés au chapitre quatre, un tel résultat est tout de même en accord avec notre propre observation d'une forte abondance d'inversions à très courte distance (<50 bp). Il est néanmoins possible que les réarrangements de type U-turn ne soient pas confinés à des répétitions inversées distantes de moins de 50 nucléotides. En effet, près de 80 % de tous les réarrangements se produisant à courte distance (<1000 bp) dans le chloroplaste des plantes de type sauvage d'*Arabidopsis* sont des inversions.

Chez la levure, l'étude de réarrangements génomiques semblables a établi que ce type de fusion entre des répétitions inversées résultait d'un mécanisme associé à la réplication, mais indépendant de la formation d'un DSB [297,298]. Tout comme le modèle que nous suggérons pour la formation de ces réarrangements U-turn, les auteurs d'une de ces études font

l'hypothèse que la formation de ce type de réarrangements serait initiée, lors du processus de réplication, par l'appariement du brin continu sur le brin matrice opposé [298]. De telles inversions génomiques sont d'ailleurs générées chez la levure par des répétitions aussi courtes que 20 nucléotides, suggérant que des microhomologies inversées peuvent mener à ce type de réarrangements [298]. De telles observations chez la levure renforcent l'idée qu'un mécanisme associé à la réplication et indépendant des mécanismes de recombinaison classiques soit à l'origine des réarrangements U-turn observés dans notre étude. De plus, il est très intéressant de constater que la mutation *rad51* cause une accumulation très importante de ce type d'inversions génomiques chez *Saccharomyces cerevisiae* [298]. Ceci est encore une fois en accord avec nos propres observations, selon lesquelles la machinerie de recombinaison limiterait l'accumulation de ce type de réarrangements génomiques.

Une étude récente a toutefois rapporté un comportement bien étrange des fourches de réplication lors de leur rencontre avec des séquences répétées et inversées [299]. En effet, il semble que les fourches de réplication bloquées, mais redémarrées adéquatement sur le bon brin matrice, resteraient hautement instables et susceptibles aux inversions. De telles fourches auraient tendances à effectuer un « U-turn » sur le mauvais brin matrice lors d'une rencontre avec des séquences répétées et inversées. Pour le moment, il nous est toutefois impossible de déterminer si un tel comportement des fourches de réplication contribue aux réarrangements de type U-turn observés dans les organites d'*Arabidopsis* et d'humain. Ce modèle alternatif de la formation de très courtes inversions reste néanmoins très similaire au modèle proposé au chapitre quatre.

5.3.3 Rôles probables de RECA1 et des protéines Whirly dans le redémarrage des fourches de réplication

La très forte accumulation de réarrangements de type U-turn chez *why1why3reca1* suggère que les protéines RECA1, WHY1 et WHY3 participent directement à la stabilisation et à la progression des fourches de réplication. Bien que notre étude n'explique pas le mécanisme par lequel ces protéines contribuent à la réplication de l'ADN et à la répression des réarrangements de type U-turn, les mécanismes bactériens de redémarrage des fourches de

réplication nous permettent d'émettre des hypothèses quant à la fonction des protéines RECA1, WHY1 et WHY3 au sein des fourches de réplication chloroplastiques.

Plusieurs études ont démontré que RecA joue un rôle central dans la progression des fourches de réplication chez *E. coli*. En effet, il a été observé par des approches génétiques et *in vitro* que RecA catalyse la régression des fourches de réplication bloquées [291,292,329]. Quoique le mécanisme par lequel RecA catalyse cette régression reste controversé, il a été suggéré que la multimérisation de RecA sur une région d'ADNss induit une torsion de la molécule d'ADN, et ce de façon dépendante de l'hydrolyse de l'ATP [revu dans 330,331]. Dans le cas d'une fourche de réplication bloquée sur le brin continu, RecA multimériserait sur la région d'ADNss en aval de l'ADN polymérase immobilisée et, par cette torsion, induirait la régression de la fourche [330]. Il apparaît probable que RECA1 possède un rôle similaire dans le chloroplaste. D'ailleurs, tout comme chez *E. coli*, la fonction de RECA1 semble dépendante de l'hydrolyse de l'ATP [208]. Il est donc vraisemblable que dans le chloroplaste, RECA1 catalyse la régression des fourches de réplication bloquées et, ce faisant, contribue à limiter le redémarrage inadéquat de la réplication par un mécanisme de type U-turn. Chez *Arabidopsis*, une étude récente a d'ailleurs démontré que la mutation du gène *RECA1* provoque une hypersensibilité au méthylviologène, un herbicide générant du superoxide chloroplastique [332]. Cette hypersensibilité supporte l'idée d'une participation de RECA1 dans la réplication de l'ADNpt, les ROS pouvant induire des lésions des bases azotées bloquant la progression des ADN polymérases [151,152].

D'autre part, des niveaux importants de réarrangements de type U-turn sont observés chez les mutants *why1why3pollb* et *why1why3reca1*, mais pas chez les simples mutants *pollb* et *reca1*. Ces résultats démontrent le rôle très important que jouent les protéines Whirly dans la répression de ce type d'instabilité génomique. La protéine de liaison à l'ADNss, SSB, stabilise l'ADNss déroulé dans les fourches de réplication chez les procaryotes. Dans les chloroplastes, la principale activité de liaison à l'ADNss semble provenir des protéines Whirly [27]. Il semble donc probable que, tout comme SSB, les protéines Whirly stabilisent l'ADNss

dans les fourches de réplication. Une telle stabilisation de l'ADNss préviendrait l'appariement entre le brin continu nouvellement synthétisé et le brin matrice opposé.

Par ailleurs, plusieurs évidences suggèrent que la protéine SSB joue un rôle central dans le processus de régression des fourches bloquées en recrutant la machinerie protéique nécessaire, dont RecA [197,333-335]. Chez SSB, ce recrutement de nombreuses protéines est accompli par l'extrémité C-terminale de la protéine, celle-ci étant hautement flexible et devenant accessible lors de liaison de la protéine à l'ADNss [revu dans 216]. Curieusement, les extrémités C-terminale des protéines Whirly sont, tout comme chez SSB, très flexibles et désordonnées [29,336]. De plus, l'analyse de la structure cristallographique des protéines Whirly suggère que cette extrémité C-terminale pourrait être relâchée lors de la liaison à l'ADNss [29]. Il semble donc possible qu'à l'instar de SSB, les protéines Whirly orchestrent elles aussi les processus du métabolisme de l'ADNpt, dont le recrutement de la machinerie assurant le redémarrage des fourches de réplication bloquées. Toutefois, puisque aucun interacteur physique des Whirly n'a été identifié à ce jour, cette question reste ouverte pour le moment.

5.3.4 Les réarrangements de type U-turn semblent être une caractéristique intrinsèque de l'ADN des organites : les implications d'une telle structure

Les réarrangements de type U-turn constituent une grande proportion des réarrangements génomiques totaux dans les organites, soit environ 25% dans le chloroplaste d'*Arabidopsis*, 38% dans la mitochondrie d'*Arabidopsis*, et 60 % dans la mitochondrie humaine (Fig. 37, page 123). Une telle abondance de ces très courtes inversions (<50 bp) soulève la possibilité que ce type de réarrangements constitue une caractéristique intrinsèque de l'ADN des organites. Suite à la ré-initiation de la réplication de l'ADN sur le brin matrice opposé, il est vraisemblable que l'extrémité 3' du brin continu nouvellement synthétisé soit liguée à l'extrémité 5' du brin discontinu (Fig. 55). Il a d'ailleurs été suggéré, chez la levure, que la formation de ce type d'inversions génomiques nécessite l'ADN polymérase delta du brin discontinu afin d'assurer l'élongation du brin continu jusqu'au fragment d'Okazaki adjacent [337]. Lors de la formation des réarrangements de type U-turn dans l'ADN des organites,

cette ligation au brin discontinu pourrait donc former une jonction d'ADN à trois brins (Fig. 55).

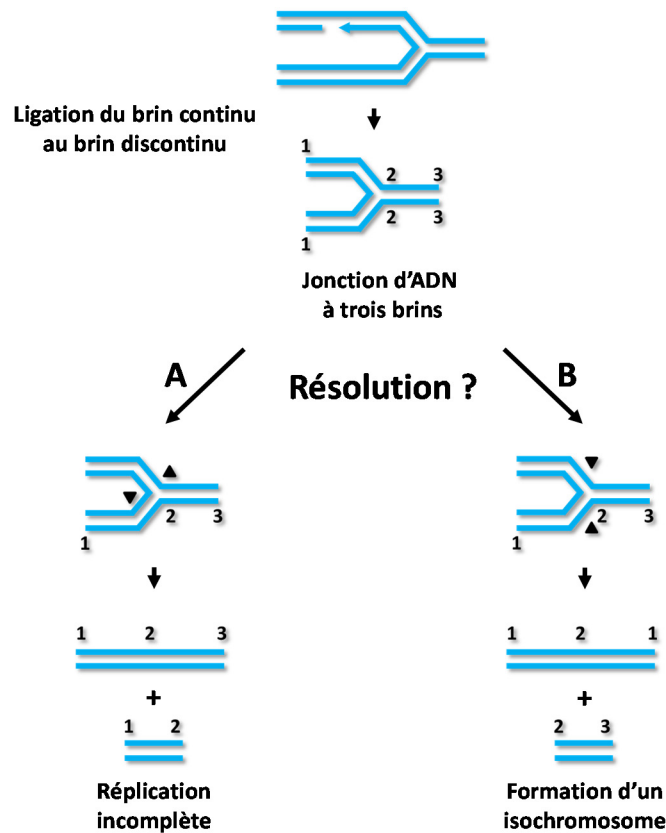


Figure 55. Conséquences possibles de la ligation du brin continu au brin discontinu lors d'un réarrangement de type U-turn. Lors d'un réarrangement de type U-turn, la ré-initiation de la synthèse du brin continu sur le brin matrice opposé pourrait mener à la ligation de l'extrémité 3' du brin continu à l'extrémité 5' du brin discontinu. Une telle ligation mènerait à la formation d'une jonction d'ADN à trois brins. La résolution d'une telle structure peut mener à deux types de produits distincts : 1) le duplex initial et une molécule d'ADN partiellement répliquée, ou 2) une molécule d'ADN palindromique (isochromosome) en plus d'une molécule d'ADN partiellement répliquée. Les nombres 1, 2 et 3 représentent des positions relatives sur la molécule d'ADN.

Une telle jonction d'ADN à trois brins implique deux possibilités : la formation d'un branchement sur la molécule d'ADN ou la résolution de la jonction d'ADN à trois brins (Fig. 55). La résolution d'une telle jonction dans l'ADN semble possible, certaines résolvases possédant la capacité de cliver une jonction d'ADN à trois brins [338,339]. Selon l'orientation

d'une telle résolution, deux types de molécules d'ADN résultantes peuvent être générées. Selon l'orientation A sur la Figure 55, la résolution engendrera un duplex d'ADN identique au duplex matrice initial et une molécule d'ADN partiellement répliquée se terminant par un DSB à une seule extrémité. Bien qu'impliquant un DSB, les conséquences d'une telle résolution restent néanmoins modestes, la réplication de la molécule inachevée pouvant être redémarrée par le processus de BIR. Toutefois, une résolution selon l'orientation B risque d'occasionner des conséquences bien plus importantes pour le maintien de la stabilité génomique. En effet, une telle résolution générera plutôt un isochromosome, c'est-à-dire un chromosome palindromique, en plus d'une molécule d'ADN brisée par un DSB à une seule extrémité.

Les brins complémentaires des réarrangements de type U-turn sont rarement observés dans nos données de séquençage. Il semble donc peu probable que de telles jonctions d'ADN à trois brins soient résolues de manière à former des isochromosomes. Conséquemment, il semble davantage probable que ces jonctions d'ADN à trois brins persistent et forment des branchements au sein de l'ADN des organites. L'ADNpt est d'ailleurs composé d'un mélange de molécules circulaires, linéaires et branchées [35]. À l'instar de l'ADNpt, des branchements sont observés dans l'ADN mitochondrial de la levure *Candida albicans* et, vraisemblablement, également chez *S. pombe* et *S. cerevisiae* [340,341]. Jusqu'à présent, ce type de branchements a été associé à l'initiation de la réplication par le processus de BIR [35,340,341]. Toutefois, en plus des événements d'initiation de la réplication par BIR, il est possible que les réarrangements de type U-turn contribuent à cette particularité de l'ADN des organites. Les mécanismes de U-turn pourraient d'ailleurs servir d'alternatives dans les organites pour redémarrer la réplication lorsque les mécanismes adéquats de redémarrage ont échoués, les conséquences de ce type de réarrangements étant bien moins graves au sein d'un génome polyploïde.

Il est intéressant de constater que l'ADN mitochondrial du cœur et du cerveau humain présente lui aussi une structure hautement branchée [342] (Fig. 56). Il a été suggéré que, dans le cœur, la grande activité de la chaîne de transport des électrons provoque une

importante production de ROS qui, à son tour, endommage l'ADN et génère des DSBs. La réparation de ces DSBs par BIR mènerait subséquemment aux branchements [343]. Il est d'ailleurs connu que les ROS initient des événements de BIR dans les mitochondries de levures [344].

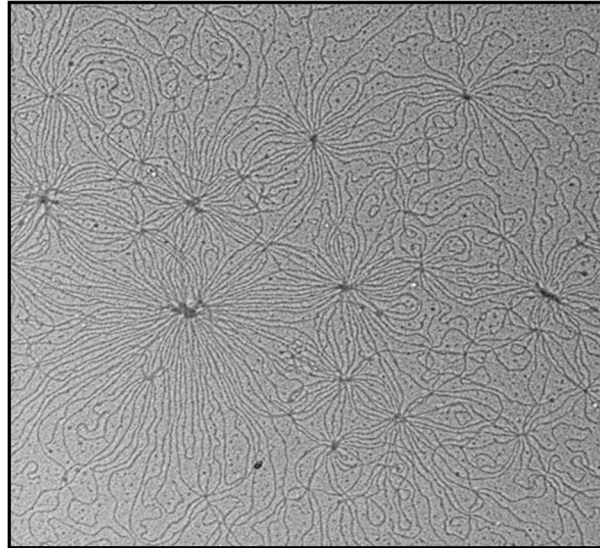


Figure 56. L'ADN mitochondrial du cœur humain est organisé en structures hautement branchées. ADN mitochondrial du cœur humain observé par microscopie électronique. Cette figure est adaptée de [343].

Néanmoins, puisque certains produits de l'oxydation des bases azotées bloquent la progression des ADN polymérases, la génération de réarrangements de type U-turn pourrait aussi contribuer à ces branchements. Cette hypothèse est de plus supportée par une étude des réarrangements génomiques mitochondriaux dans le cœur de souris [345]. Curieusement, bien que l'ADN mitochondrial du cœur de souris ne soit que peu branché par rapport au cœur humain, des branchements similaires peuvent être induits par la surexpression de l'hélicase mitochondriale Twinkle [346]. Dans de telles conditions, l'analyse de l'instabilité génomique mitochondriale par séquençage de nouvelle génération révèle une forte accumulation de réarrangements génomiques se produisant à courte distance [345]. Aucune analyse des types de réarrangements génomiques n'a toutefois été effectuée. La

contribution des réarrangements de type U-turn aux branchements de l'ADN mitochondrial reste donc inconnue pour le moment.

5.3.5 Les conséquences possibles d'une accumulation excessive de réarrangements de type U-turn dans l'ADN chloroplastique

Alors que les triples mutants *why1why3pollb* et *why1why3reca1* présentent des patrons d'instabilité génomique extrêmement similaires à grande distance (>1000 bp), une accumulation bien plus importante de réarrangements à courte distance (<1000 bp) est observée chez *why1why3reca1* (Fig. 39, page 129). Ces deux lignées diffèrent aussi par leurs phénotypes, seul *why1why3reca1* présentant une intense variégation blanche. Il apparaît donc probable que cette importante accumulation de réarrangements génomiques à courte distance soit responsable de l'apparition de ce très sévère phénotype. Parmi ces courts réarrangements (<1000 bp), ceux de type U-turn (<50 bp) sont omniprésents chez *why1why3reca1* et représentent environ 60 % de toute l'instabilité génomique à courte distance chez cette lignée (<1000 bp). Ce chiffre est d'ailleurs probablement sous-estimé, les réarrangements de type U-turn n'étant vraisemblablement pas confinés à une distance maximale de 50 bp. En effet, plus de 94 % des réarrangements à courte distance (<1000 bp) sont des inversions chez *why1why3reca1*. Ceci suggère qu'une importante accumulation de réarrangements de type U-turn altère le développement normal de la plante et se manifeste par l'apparition d'un phénotype de variégation blanche.

Une telle corrélation soulève toutefois la question des impacts moléculaires que peut causer une accumulation excessive de réarrangements de type U-turn. À priori, il peut sembler logique que ce type de réarrangements génomiques ne fassent qu'aggraver le déséquilibre photosynthétique induit par les autres types d'instabilité génomique. Ceci semble néanmoins peu probable, les triples mutants *why1why3pollb* et *why1why3reca1* présentant une efficacité photosynthétique similaire et le phénotype de variégation blanche ne pouvant être modulé par l'intensité lumineuse (Fig. 34 et 35, pages 108 et 111). Vraisemblablement, ce type d'instabilité génomique induit donc d'autres effets délétères dans le chloroplaste.

Tel que discuté dans la section précédente, il est possible que les réarrangements de type U-turn forment des branchements au sein de l'ADNpt. Il est donc envisageable que les nucléoïdes du plastide se retrouvent hautement entremêlés lors d'une accumulation excessive de ce type de réarrangements et que de tels branchements empêchent la biogénèse des chloroplastes. La différenciation des proplastés en chloroplastes est étroitement liée à la formation des membranes thylakoïdes et de la PET [347,348]. La mutation d'une composante essentielle de la PET, *FtsH2*, mène d'ailleurs à l'apparition de variégation blanche associée à des cellules viables, mais dans lesquelles la différenciation des proplastés en chloroplastes est inhibée [347]. De manière similaire, il est possible que de nombreux branchements dans l'ADNpt de *why1why3reca1* affectent gravement l'expression des gènes photosynthétiques et empêchent la formation normale de la PET et des membranes thylakoïdes. Une telle éventualité rendrait la différenciation des proplastés impossible et expliquerait la formation de secteurs blancs viables sur les tissus de *why1why3reca1* (Fig. 33, page 106).

Par ailleurs, la différenciation des proplastés en plastides spécialisés ne suffit pas à générer un nombre très élevé de chloroplastes dans les cellules végétales. Cette différenciation est complétée par un processus de scission des chloroplastes matures. Lors de ce processus, les nucléoïdes sont séparés en quantité presque égale entre les deux chloroplastes fils [39,40]. En plus d'affecter la différenciation des proplastés en chloroplastes, il est possible que de nombreux branchements dans l'ADNpt altèrent la ségrégation normale des nucléoïdes lors de la division des chloroplastes. Deux cas potentiellement délétères sont envisageables : 1) un des chloroplastes fils hérite d'une trop grande quantité d'ADNpt hautement branché et probablement non-fonctionnel et 2) un des chloroplastes fils hérite d'une quantité trop faible d'ADN. Ces deux cas pourraient d'ailleurs être à l'origine des secteurs variégués blancs chez *why1why3reca1*. Une transmission d'ADNpt hautement branché et non-fonctionnel pourrait aussi se refléter lors de la transmission maternelle de l'ADNpt à l'embryon et expliquer le haut pourcentage de létalité des graines chez *why1why3reca1* (Fig. 44, page 145). De même, il est possible qu'une transmission trop faible d'ADNpt mène à la létalité de l'embryon. Une

telle éventualité n'a toutefois pas été testée pour le moment. Curieusement, aucune transmission maternelle du phénotype de variégation blanche n'est observée lors de croisements réciproques entre une plante *why1why3reca1* et une plante de type sauvage (E. Zampini, résultats préliminaires non-publiés). Néanmoins, un tel croisement (♀ *why1why3reca1*; ♂ WT) mène également à un faible taux de germination et appuie l'éventualité d'une ségrégation inadéquate de l'ADNpt chez *why1why3reca1* (E. Zampini, résultats préliminaires non-publiés). Cette hypothèse est également supportée par une étude distincte démontrant le rôle des ADN gyrases chloroplastiques dans la décaténation, ou séparation, des chromosomes chloroplastiques nouvellement répliqués [71]. En effet, chez des lignées de *N. benthamiana* ayant une expression réduite de chacune des sous-unités des gyrases, les nucléoïdes chloroplastiques sont anormalement élargis et paraissent hautement entremêlés. De plus, chez ces lignées, alors que certains chloroplastes contiennent une quantité d'ADNpt bien plus importante que les plantes de type sauvage, d'autres en contiennent une quantité bien plus faible. Cette étude supporte donc l'idée que l'entremêlement de l'ADNpt affecte la ségrégation normale des nucléoïdes chloroplastiques. Il est aussi intéressant de constater que, comme chez *why1why3reca1*, un phénotype de variégation et de morphologie anormale est observé chez les lignées ayant subi une déplétion en gyrases.

5.3.6 Perspectives et expériences futures

Le mécanisme de formation des réarrangements de type U-turn proposé au chapitre quatre soulève de nombreuses questions, tant pour la validité du modèle proposé que pour ses implications. Tel que mentionné plus tôt dans ce chapitre, ce modèle implique qu'une augmentation de la fréquence des fourches de réplication bloquées devrait accroître l'accumulation de réarrangements de type U-turn. Chez *S. pombe*, un traitement au méthanesulfonate de méthyle (MMS), un agent alkylant les bases azotées de l'ADN, induit d'ailleurs une accumulation d'inversions génomiques associées à la réplication et indépendantes de la formation d'un DSB [297]. Une croissance de plants d'*Arabidopsis* en présence d'agents induisant des lésions dans l'ADN permettrait de tester aisément la

dépendance entre la formation des réarrangements de type U-turn et l'inhibition de la progression des fourches de réplication [349]. Ainsi, il pourrait être vérifié si une exposition aux rayons UV-B ou au MMS aggrave l'accumulation de réarrangements de type U-turn. Il serait également intéressant de tester des herbicides générant des ROS chloroplastiques, tels que le méthylviologène et l'atrazine [255,256]. Curieusement, même à des doses inhibant la croissance de la plante, ces deux herbicides ne semblent induire aucune augmentation des réarrangements chloroplastiques de type MMBIR (Fig. 29, page 91). Ceci n'exclut cependant pas la possibilité que ces composés induisent une accumulation de réarrangements de type U-turn. La mutation *reca1* induit d'ailleurs une hypersensibilité au méthylviologène chez *Arabidopsis* [332]. D'autre part, le 3'-azidothymidine (AZT) est un analogue de nucléoside couramment utilisé pour le traitement du VIH. Durant le processus de transcription inverse par le virus, l'incorporation de l'AZT dans l'ADN bloque la progression de la polymérisation [revu dans 350]. Cette molécule est toutefois associée à une toxicité mitochondriale, probablement en raison de son incorporation dans l'ADN de cet organelle par l'ADN polymérase γ [350]. De plus, durant le processus de réplication de l'ADN bactérien, l'AZT induit des mutations et des réarrangements génomiques associés à des répétitions inversées [301,351]. Il apparaît donc possible que la toxicité mitochondriale associée à l'AZT soit liée à une accumulation de réarrangements de type U-turn.

Chez la bactérie, la régression des fourches de réplication bloquées n'est pas catalysée que par la recombinaise RecA. En effet, RecG, une hélicase s'associant aux structures d'ADN branchées, catalyse la migration des jonctions d'ADN et permet, de cette façon, la régression des fourches de réplication bloquées [352,353]. RecG assure en plus le ré-enroulement des fourches de réplication afin de redémarrer la synthèse d'ADN [354]. De manière intéressante, il a été démontré qu'un homologue de cette ADN hélicase bactérienne participe au maintien de l'intégrité génomique dans le plastide et dans la mitochondrie de la mousse *Physcomitrella patens* [224]. La mutation *recg* cause d'ailleurs une accumulation de réarrangements génomiques dépendants de microhomologies dans les organites de cet organisme. Il est tentant de spéculer que cette instabilité génomique pourrait être associée à

l'incapacité du mutant *recg* à redémarrer efficacement les fourches de réplication bloquées et à la formation conséquente de DSBs. Un homologue de cette hélicase est d'ailleurs encodé chez *Arabidopsis* et possède plus de 33% d'identité avec son équivalent bactérien (E. Zampini, résultat préliminaire). Plusieurs outils de prédiction de localisation intracellulaire suggèrent en plus que RECG serait ciblé au plastide suite à son expression (E. Zampini, résultat préliminaire). De plus, certaines lignées mutantes de *RECG* semblent disponibles chez *Arabidopsis* et pourraient constituer un outil précieux pour l'étude de ce gène par une approche génétique. Afin de vérifier l'implication de *RECG* dans le redémarrage des fourches de réplication bloquées, des lignées mutantes *recg* pourraient, entre autres, être utilisées pour vérifier si des interactions génétiques existent chez des lignées *why1why3recg* et *reca1recg*.

Finalement, il semble possible que l'importante accumulation de réarrangements de type U-turn chez *why1why3reca1* soit associée à de nombreux branchements au sein de l'ADNpt qui, à leur tour, provoquent une ségrégation anormale des nucléoïdes lors de la division chloroplastique. De manière à vérifier si le phénotype de variégation blanche observé chez *why1why3reca1* est associé à l'entremêlement de l'ADNpt plutôt qu'à l'altération de la séquence d'ADN par les réarrangements de type U-turn, une résolvasse clivant les jonctions d'ADN à trois brins pourrait être dirigée au chloroplaste [338,339]. Il est possible que le phénotype de variégation blanche soit complémenté chez une telle plante transgénique. D'autre part, si les réarrangements de type U-turn mènent à une ségrégation anormale de l'ADNpt chez *why1why3reca1*, il est probable que des populations distinctes de chloroplastes coexistent au sein des tissus. L'analyse de plantes entières rend toutefois impossible la comparaison de l'ADNpt entre des tissus verts et des tissus variégés blancs. Afin de pallier un tel handicap, une approche par cytométrie en flux pourrait être employée de manière à séparer les chloroplastes dépourvus de chlorophylle de ceux d'apparence normale [355]. L'ADNpt de ces chloroplastes pourrait également être coloré au 4',6'-diamidino-2-phénylindole (DAPI) afin de trier les populations de chloroplastes blancs ou verts selon leur quantité d'ADN. Une telle analyse permettrait de visualiser les différentes populations de

chloroplastes coexistant au sein de la lignée *why1why3reca1* et d'établir si des niveaux importants de réarrangements de type U-turn corréleraient avec une ségrégation anormale de l'ADNpt. Enfin, l'ADNpt de ces différentes populations de chloroplastes pourrait être purifié et visualisé par microscopie électronique afin d'établir la quantité de branchements au sein des nucléoïdes. De plus, les niveaux de réarrangements de type U-turn pourraient être analysés par séquençage Illumina au sein de chacune de ces populations chloroplastiques. Une telle méthode permettrait donc, possiblement, d'établir une corrélation entre la quantité de réarrangements de type U-turn dans l'ADNpt, l'abondance de branchements au sein des nucléoïdes, la ségrégation de l'ADNpt et la fonction chloroplastique.

Conclusion

Les résultats obtenus au cours de mes études doctorales ont permis de mieux comprendre certaines des principales causes d'instabilité génomique dans les organites et quelques-unes de leurs conséquences physiologiques et moléculaires. Tout comme de nombreuses autres études, ces résultats illustrent les conséquences contrastantes d'un déséquilibre rédox associé au chloroplaste, tant pour ses effets délétères que pour son important rôle dans la signalisation cellulaire et dans l'adaptation à des conditions de stress. Ce projet réalisé au laboratoire s'inscrit dans l'un des principaux défis auquel fait face la biologie végétale actuellement. Afin de pouvoir répondre aux demandes sans cesse croissantes de productivité agricole, il est en effet essentiel d'approfondir notre compréhension des impacts des stress abiotiques et des conditions induisant un déséquilibre photosynthétique chez les plantes.

D'autre part, la nature polyploïde de l'ADN des organites et la grande difficulté associée à la manipulation génétique de ces génomes rendent ardue l'étude du métabolisme de leur ADN. L'approche bio-informatique présentée dans ce travail pourrait toutefois constituer un outil précieux facilitant l'étude de l'instabilité génomique des organites. En effet, celle-ci permettrait d'évaluer, entre autres, si certains patrons particuliers d'instabilité génomique constituent une caractéristique distinctive de diverses pathologies possédant une composante mitochondriale. De plus, cette approche a révélé que les réarrangements de type U-turn semblent constituer une caractéristique intrinsèque de l'ADN des organites, et ce, dans deux règnes distincts. Dans le chloroplaste d'*Arabidopsis*, des niveaux importants de ce type d'instabilité génomique semblent être associés à un sévère trouble développemental de la plante. Il convient donc de se questionner sur les conséquences que pourrait occasionner une accumulation excessive de ce type de réarrangements génomiques dans les mitochondries humaines. Finalement, la combinaison de cette méthode d'analyse et d'une approche génétique a permis de déterminer que la recombinaison chloroplastique RECA1 et les protéines de liaison à l'ADN simple-brin nommées Whirly maintiennent la stabilité génomique chloroplastique en empêchant la formation des réarrangements de type U-turn. Chez l'humain, la recombinaison RAD51 et la protéine de liaison à l'ADN simple-brin mtSSB

participent à la réplication de l'ADN mitochondrial [356,357]. Notre étude soulève donc la possibilité que des mécanismes similaires assurent une réplication adéquate et efficace de l'ADN mitochondrial humain. De telles éventualités pourront être explorées dans le futur.

Références

1. O'Connor CM, Adams JU (2010) Essentials of Cell Biology. Cambridge, MA: NPG Education.
2. Wise R (2006) The Diversity of Plastid Form and Function. In: Wise RR, Hooper JK, editors. The Structure and Function of Plastids. Netherlands: Springer. pp. 3-26.
3. Lopez-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. *International Journal of Developmental Biology* 49: 557-577.
4. Martin W, Kowallik KV (1999) Annotated English translation of Mereschowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *European Journal of Phycology* 34: 287-295.
5. Sato N (2007) Origin and Evolution of Plastids: Genomic View on the Unification and Diversity of Plastids. Springer: 75-102.
6. McFadden GI (2014) Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb Perspect Biol* 6: a016105.
7. Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. *Science* 283: 1476-1481.
8. Douglas SE, Turner S (1991) Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. *Journal of Molecular Evolution* 33: 267-273.
9. Nelissen B, Van de Peer Y, Wilmotte A, De Wachter R (1995) An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. *Molecular Biology and Evolution* 12: 1166-1173.
10. Ochoa de Alda JA, Esteban R, Diago ML, Houmard J (2014) The plastid ancestor originated among one of the major cyanobacterial lineages. *Nat Commun* 5: 4937.
11. Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annual Review of Plant Biology* 60: 115-138.
12. Bock R, Timmis JN (2008) Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays* 30: 556-566.
13. Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, et al. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393: 162-165.
14. Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422: 72-76.
15. Stegemann S, Hartmann S, Ruf S, Bock R (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 100: 8828-8833.
16. Jarvis P, Lopez-Juez E (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nature Reviews: Molecular Cell Biology* 14: 787-802.
17. Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and Why? *Plant physiology* 118: 9-17.
18. Schweer J (2010) Plant sigma factors come of age: flexible transcription factor net-work for regulated plastid gene expression. *Endocytobiosis Cell Res* 20: 1-20.
19. Muller HJ (1964) The Relation of Recombination to Mutational Advance. *Mutation research* 106: 2-9.
20. Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* 84: 9054-9058.
21. Greiner S, Sobanski J, Bock R (2015) Why are most organelle genomes transmitted maternally? *Bioessays* 37: 80-94.

22. Khakhlova O, Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant Journal* 46: 85-94.
23. Green BR (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant Journal* 66: 34-44.
24. Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, et al. (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* 14: 2659-2679.
25. Yamane K, Yano K, Kawahara T (2006) Pattern and rate of indel evolution inferred from whole chloroplast intergenic regions in sugarcane, maize and rice. *DNA Res* 13: 197-204.
26. Kawata M, Harada T, Shimamoto Y, Oono K, Takaiwa F (1997) Short inverted repeats function as hotspots of intermolecular recombination giving rise to oligomers of deleted plastid DNAs (ptDNAs). *Current Genetics* 31: 179-184.
27. Maréchal A, Parent JS, Veronneau-Lafortune F, Joyeux A, Lang BF, et al. (2009) Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 106: 14693-14698.
28. Kwon T, Huq E, Herrin DL (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 107: 13954-13959.
29. Cappadocia L, Maréchal A, Parent JS, Lepage E, Sygusch J, et al. (2010) Crystal structures of DNA-Whirly complexes and their role in *Arabidopsis* organelle genome repair. *Plant Cell* 22: 1849-1867.
30. Bonham-Smith PC, Bourque DP (1989) Translation of chloroplast-encoded mRNA: potential initiation and termination signals. *Nucleic Acids Research* 17: 2057-2080.
31. Hirose T, Sugiura M (2004) Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. *Plant and Cell Physiology* 45: 114-117.
32. Lohse M, Drechsel O, Kahlau S, Bock R (2013) OrganellarGenomeDRAW--a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* 41: W575-581.
33. Zoschke R, Liere K, Borner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant Journal* 50: 710-722.
34. Palmer JD (1983) Chloroplast DNA Exists in 2 Orientations. *Nature* 301: 92-93.
35. Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16: 1661-1666.
36. Deng XW, Wing RA, Gruissem W (1989) The chloroplast genome exists in multimeric forms. *Proceedings of the National Academy of Sciences of the United States of America* 86: 4156-4160.
37. Lilly JW, Havey MJ, Jackson SA, Jiang J (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *Plant Cell* 13: 245-254.
38. Sakamoto W, Miyagishima SY, Jarvis P (2008) Chloroplast biogenesis: control of plastid development, protein import, division and inheritance. *Arabidopsis Book* 6: e0110.
39. Kuroiwa T, Suzuki T, Ogawa K, Kawano S (1981) The Chloroplast Nucleus: Distribution, Number, Size, and Shape, and a Model for the Multiplication of the Chloroplast Genome during Chloroplast Development. *Plant & Cell Physiology* 22: 381-396.
40. Hashimoto H, Possingham JV (1989) Effect of light on the chloroplast division cycle and DNA synthesis in cultured leaf discs of spinach. *Plant Physiology* 89: 1178-1183.
41. Pfalz J, Pfannschmidt T (2013) Essential nucleoid proteins in early chloroplast development. *Trends in Plant Science* 18: 186-194.

42. Majeran W, Friso G, Asakura Y, Qu X, Huang M, et al. (2012) Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiology* 158: 156-189.
43. Parent JS, Lepage E, Brisson N (2011) Divergent roles for the two PolI-like organelle DNA polymerases of *Arabidopsis*. *Plant Physiology* 156: 254-262.
44. Cerutti H, Ibrahim HZ, Jagendorf AT (1993) Treatment of pea (*Pisum sativum* L.) protoplasts with DNA-damaging agents induces a 39-kilodalton chloroplast protein immunologically related to *Escherichia coli* RecA. *Plant Physiology* 102: 155-163.
45. Wall MK, Mitchenall LA, Maxwell A (2004) *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 101: 7821-7826.
46. Krupinska K, Oetke S, Desel C, Mulisch M, Schafer A, et al. (2014) WHIRLY1 is a major organizer of chloroplast nucleoids. *Front Plant Sci* 5: 432.
47. Schroter Y, Steiner S, Matthai K, Pfannschmidt T (2010) Analysis of oligomeric protein complexes in the chloroplast sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. *Proteomics* 10: 2191-2204.
48. Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmüller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18: 176-197.
49. Serino G, Maliga P (1998) RNA polymerase subunits encoded by the plastid *rpo* genes are not shared with the nucleus-encoded plastid enzyme. *Plant Physiology* 117: 1165-1170.
50. Chang CC, Sheen J, Bligny M, Niwa Y, Lerbs-Mache S, et al. (1999) Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. *Plant Cell* 11: 911-926.
51. Hedtke B, Börner T, Weihe A (2000) One RNA polymerase serving two genomes. *EMBO Rep* 1: 435-440.
52. Zhelyazkova P, Sharma CM, Forstner KU, Liere K, Vogel J, et al. (2012) The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *Plant Cell* 24: 123-136.
53. Drechsel O, Bock R (2011) Selection of Shine-Dalgarno sequences in plastids. *Nucleic Acids Research* 39: 1427-1438.
54. Tiller N, Bock R (2014) The translational apparatus of plastids and its role in plant development. *Mol Plant* 7: 1105-1120.
55. Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* 65: 415-442.
56. Beligni MV, Yamaguchi K, Mayfield SP (2004) Chloroplast elongation factor ts pro-protein is an evolutionarily conserved fusion with the s1 domain-containing plastid-specific ribosomal protein-7. *Plant Cell* 16: 3357-3369.
57. Albrecht V, Ingenfeld A, Apel K (2006) Characterization of the snowy cotyledon 1 mutant of *Arabidopsis thaliana*: the impact of chloroplast elongation factor G on chloroplast development and plant vitality. *Plant Molecular Biology* 60: 507-518.
58. Ruppel NJ, Hangarter RP (2007) Mutations in a plastid-localized elongation factor G alter early stages of plastid development in *Arabidopsis thaliana*. *BMC Plant Biology* 7: 37.
59. Shen Y, Li C, McCarty DR, Meeley R, Tan BC (2013) Embryo defective12 encodes the plastid initiation factor 3 and is essential for embryogenesis in maize. *Plant Journal* 74: 792-804.
60. Alkatib S, Scharff LB, Rogalski M, Fleischmann TT, Matthes A, et al. (2012) The contributions of wobbling and superwobbling to the reading of the genetic code. *PLoS Genetics* 8: e1003076.

61. Draper CK, Hays JB (2000) Replication of chloroplast, mitochondrial and nuclear DNA during growth of unirradiated and UVB-irradiated Arabidopsis leaves. *Plant Journal* 23: 255-265.
62. Sala F, Amileni AR, Parisi B, Spadari S (1980) A γ -like DNA Polymerase in Spinach Chloroplasts. *Eur J Biochem* 112: 211-217.
63. Kimura S, Uchiyama Y, Kasai N, Namekawa S, Saotome A, et al. (2002) A novel DNA polymerase homologous to Escherichia coli DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). *Nucleic Acids Research* 30: 1585-1592.
64. Mori Y, Kimura S, Saotome A, Kasai N, Sakaguchi N, et al. (2005) Plastid DNA polymerases from higher plants, Arabidopsis thaliana. *Biochemical and Biophysical Research Communications* 334: 43-50.
65. Elo A, Lyznik A, Gonzalez DO, Kachman SD, Mackenzie SA (2003) Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the Arabidopsis genome. *Plant Cell* 15: 1619-1631.
66. Ono Y, Sakai A, Takechi K, Takio S, Takusagawa M, et al. (2007) NtPoll-like1 and NtPoll-like2, bacterial DNA polymerase I homologs isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. *Plant and Cell Physiology* 48: 1679-1692.
67. Nielsen BL, Rajasekhar VK, Tewari KK (1991) Pea chloroplast DNA primase: characterization and role in initiation of replication. *Plant Molecular Biology* 16: 1019-1034.
68. Tuteja N, Phan TN (1998) A chloroplast DNA helicase II from pea that prefers fork-like replication structures. *Plant Physiology* 118: 1029-1038.
69. Lam E, Chua NH (1987) Chloroplast DNA gyrase and in vitro regulation of transcription by template topology and novobiocin. *Plant Molecular Biology* 8: 415-424.
70. Saotome A, Kimura S, Mori Y, Uchiyama Y, Morohashi K, et al. (2006) Characterization of four RecQ homologues from rice (*Oryza sativa* L. cv. Nipponbare). *Biochemical and Biophysical Research Communications* 345: 1283-1291.
71. Cho HS, Lee SS, Kim KD, Hwang I, Lim JS, et al. (2004) DNA gyrase is involved in chloroplast nucleoid partitioning. *Plant Cell* 16: 2665-2682.
72. Day A, Madesis P (2008) DNA replication, recombination, and repair in plastids. In: Bock R, editor. *Cell and Molecular Biology of Plastids*. Berlin/Heidelberg: Springer. pp. 65-119.
73. Kolodner R, Tewari KK (1975) Presence of displacement loops in the covalently closed circular chloroplast deoxyribonucleic acid from higher plants. *Journal of Biological Chemistry* 250: 8840-8847.
74. Zampini E, Truche S, Lepage E, Tremblay-Belzile S, Brisson N (2015) Somatic Genome Variation of Animals, Plants and Microorganisms, Chapitre 17 : Plastid Genome Stability and Repair. Wiley, sous presse.
75. Kolodner RD, Tewari KK (1975) Chloroplast DNA from higher plants replicates by both the Cairns and the rolling circle mechanism. *Nature* 256: 708-711.
76. Scharff LB, Koop HU (2007) Targeted inactivation of the tobacco plastome origins of replication A and B. *Plant Journal* 50: 782-794.
77. Woelfle MA, Thompson RJ, Mosig G (1993) Roles of novobiocin-sensitive topoisomerases in chloroplast DNA replication in Chlamydomonas reinhardtii. *Nucleic Acids Research* 21: 4231-4238.
78. Asada K (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiology* 141: 391-396.
79. Suzuki N, Koussevitzky S, Mittler R, Miller G (2012) ROS and redox signalling in the response of plants to abiotic stress. *Plant Cell Environ* 35: 259-270.

80. Kim C, Apel K (2013) Singlet oxygen-mediated signaling in plants: moving from flu to wild type reveals an increasing complexity. *Photosynth Res* 116: 455-464.
81. Kato Y, Miura E, Ido K, Ifuku K, Sakamoto W (2009) The variegated mutants lacking chloroplastic FtsHs are defective in D1 degradation and accumulate reactive oxygen species. *Plant Physiology* 151: 1790-1801.
82. Ruckle ME, Burgoon LD, Lawrence LA, Sinkler CA, Larkin RM (2012) Plastids are major regulators of light signaling in Arabidopsis. *Plant Physiology* 159: 366-390.
83. Muller P, Li XP, Niyogi KK (2001) Non-photochemical quenching. A response to excess light energy. *Plant Physiology* 125: 1558-1566.
84. Minagawa J (2011) State transitions--the molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochimica et Biophysica Acta* 1807: 897-905.
85. Escoubas JM, Lomas M, LaRoche J, Falkowski PG (1995) Light intensity regulation of cab gene transcription is signaled by the redox state of the plastoquinone pool. *Proceedings of the National Academy of Sciences of the United States of America* 92: 10237-10241.
86. Bellafiore S, Barneche F, Peltier G, Rochaix JD (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature* 433: 892-895.
87. Pesaresi P, Hertle A, Pribil M, Schneider A, Kleine T, et al. (2010) Optimizing photosynthesis under fluctuating light: the role of the Arabidopsis STN7 kinase. *Plant Signal Behav* 5: 21-25.
88. Ballottari M, Dall'Osto L, Morosinotto T, Bassi R (2007) Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *Journal of Biological Chemistry* 282: 8947-8958.
89. Grace SC, Logan BA (1996) Acclimation of Foliar Antioxidant Systems to Growth Irradiance in Three Broad-Leaved Evergreen Species. *Plant Physiology* 112: 1631-1640.
90. Estavillo GM, Chan KX, Phua SY, Pogson BJ (2012) Reconsidering the nature and mode of action of metabolite retrograde signals from the chloroplast. *Front Plant Sci* 3: 300.
91. Bonardi V, Pesaresi P, Becker T, Schleiff E, Wagner R, et al. (2005) Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. *Nature* 437: 1179-1182.
92. Pesaresi P, Hertle A, Pribil M, Kleine T, Wagner R, et al. (2009) Arabidopsis STN7 kinase provides a link between short- and long-term photosynthetic acclimation. *Plant Cell* 21: 2402-2423.
93. Brautigam K, Dietzel L, Kleine T, Stroher E, Wormuth D, et al. (2009) Dynamic plastid redox signals integrate gene expression and metabolism to induce distinct metabolic states in photosynthetic acclimation in Arabidopsis. *Plant Cell* 21: 2715-2732.
94. Piippo M, Allahverdiyeva Y, Paakkarinen V, Suoranta UM, Battchikova N, et al. (2006) Chloroplast-mediated regulation of nuclear genes in Arabidopsis thaliana in the absence of light stress. *Physiological Genomics* 25: 142-152.
95. Meskauskiene R, Nater M, Goslings D, Kessler F, op den Camp R, et al. (2001) FLU: a negative regulator of chlorophyll biosynthesis in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* 98: 12826-12831.
96. op den Camp RG, Przybyla D, Ochsenbein C, Laloi C, Kim C, et al. (2003) Rapid induction of distinct stress responses after the release of singlet oxygen in Arabidopsis. *Plant Cell* 15: 2320-2332.
97. Wagner D, Przybyla D, Op den Camp R, Kim C, Landgraf F, et al. (2004) The genetic basis of singlet oxygen-induced stress responses of Arabidopsis thaliana. *Science* 306: 1183-1185.
98. Lee KP, Kim C, Landgraf F, Apel K (2007) EXECUTER1- and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* 104: 10270-10275.

99. Kim C, Meskauskienė R, Zhang S, Lee KP, Lakshmanan Ashok M, et al. (2012) Chloroplasts of *Arabidopsis* are the source and a primary target of a plant-specific programmed cell death signaling pathway. *Plant Cell* 24: 3026-3039.
100. Maruta T, Noshi M, Tanouchi A, Tamoi M, Yabuta Y, et al. (2012) H₂O₂-triggered retrograde signaling from chloroplasts to nucleus plays specific role in response to stress. *Journal of Biological Chemistry* 287: 11717-11729.
101. Laloi C, Stachowiak M, Pers-Kamczyc E, Warzych E, Murgia I, et al. (2007) Cross-talk between singlet oxygen- and hydrogen peroxide-dependent signaling of stress responses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 104: 672-677.
102. Xiao Y, Savchenko T, Baidoo EE, Chehab WE, Hayden DM, et al. (2012) Retrograde signaling by the plastidial metabolite MEcPP regulates expression of nuclear stress-response genes. *Cell* 149: 1525-1535.
103. Estavillo GM, Crisp PA, Pornsiriwong W, Wirtz M, Collinge D, et al. (2011) Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in *Arabidopsis*. *Plant Cell* 23: 3992-4012.
104. Ramel F, Birtic S, Ginies C, Soubigou-Taconnat L, Triantaphylides C, et al. (2012) Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences of the United States of America* 109: 5535-5540.
105. Sun X, Feng P, Xu X, Guo H, Ma J, et al. (2011) A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. *Nat Commun* 2: 477.
106. Gadjev I, Vanderauwera S, Gechev TS, Laloi C, Minkov IN, et al. (2006) Transcriptomic footprints disclose specificity of reactive oxygen species signaling in *Arabidopsis*. *Plant Physiology* 141: 436-445.
107. Leon P, Gregorio J, Cordoba E (2012) ABI4 and its role in chloroplast retrograde communication. *Front Plant Sci* 3: 304.
108. Fitter DW, Martin DJ, Copley MJ, Scotland RW, Langdale JA (2002) GLK gene pairs regulate chloroplast development in diverse plant species. *Plant Journal* 31: 713-727.
109. Kakizaki T, Matsumura H, Nakayama K, Che FS, Terauchi R, et al. (2009) Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiology* 151: 1339-1353.
110. Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, et al. (2007) Signals from chloroplasts converge to regulate nuclear gene expression. *Science* 316: 715-719.
111. Davletova S, Schlauch K, Coutu J, Mittler R (2005) The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in *Arabidopsis*. *Plant Physiology* 139: 847-856.
112. Mittler R, Kim Y, Song L, Coutu J, Coutu A, et al. (2006) Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to abiotic stress. *FEBS Letters* 580: 6537-6542.
113. Rizhsky L, Davletova S, Liang H, Mittler R (2004) The zinc finger protein Zat12 is required for cytosolic ascorbate peroxidase 1 expression during oxidative stress in *Arabidopsis*. *Journal of Biological Chemistry* 279: 11736-11743.
114. Petrillo E, Godoy Herz MA, Fuchs A, Reifer D, Fuller J, et al. (2014) A chloroplast retrograde signal regulates nuclear alternative splicing. *Science* 344: 427-430.
115. Wilson PB, Estavillo GM, Field KJ, Pornsiriwong W, Carroll AJ, et al. (2009) The nucleotidase/phosphatase SAL1 is a negative regulator of drought tolerance in *Arabidopsis*. *Plant Journal* 58: 299-317.

116. Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, et al. (2003) Early light-induced proteins protect Arabidopsis from photooxidative stress. *Proceedings of the National Academy of Sciences of the United States of America* 100: 4921-4926.
117. Potter E, Kloppstech K (1993) Effects of light stress on the expression of early light-inducible proteins in barley. *European Journal of Biochemistry* 214: 779-786.
118. Grimm B, Kruse E, Kloppstech K (1989) Transiently expressed early light-inducible thylakoid proteins share transmembrane domains with light-harvesting chlorophyll binding proteins. *Plant Molecular Biology* 13: 583-593.
119. Green BR, Kuhlbrandt W (1995) Sequence conservation of light-harvesting and stress-response proteins in relation to the three-dimensional molecular structure of LHCII. *Photosynth Res* 44: 139-148.
120. Adamska I, Roobol-Boza M, Lindahl M, Andersson B (1999) Isolation of pigment-binding early light-inducible proteins from pea. *European Journal of Biochemistry* 260: 453-460.
121. Tzvetkova-Chevolleau T, Franck F, Alawady AE, Dall'Osto L, Carriere F, et al. (2007) The light stress-induced protein ELIP2 is a regulator of chlorophyll synthesis in *Arabidopsis thaliana*. *Plant Journal* 50: 795-809.
122. Triantaphylides C, Havaux M (2009) Singlet oxygen in plants: production, detoxification and signaling. *Trends in Plant Science* 14: 219-228.
123. Gill SS, Tuteja N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry* 48: 909-930.
124. Balasubramanian B, Pogozelski WK, Tullius TD (1998) DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proceedings of the National Academy of Sciences of the United States of America* 95: 9738-9743.
125. Halliwell B, Gutteridge JM (1984) Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochemical Journal* 219: 1-14.
126. Pilon M, Ravet K, Tapken W (2011) The biogenesis and physiological function of chloroplast superoxide dismutases. *Biochimica et Biophysica Acta* 1807: 989-998.
127. Myouga F, Hosoda C, Umezawa T, Iizumi H, Kuromori T, et al. (2008) A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in *Arabidopsis*. *Plant Cell* 20: 3148-3162.
128. Kliebenstein DJ, Monde RA, Last RL (1998) Superoxide dismutase in *Arabidopsis*: an eclectic enzyme family with disparate regulation and protein localization. *Plant Physiology* 118: 637-650.
129. Ogawa K, Kanematsu S, Takabe K, Asada K (1995) Attachment of CuZn-Superoxide Dismutase to Thylakoid Membranes at the Site of Superoxide Generation (Psi) in Spinach-Chloroplasts - Detection by Immunogold Labeling after Rapid Freezing and Substitution Method. *Plant and Cell Physiology* 36: 565-573.
130. Rizhsky L, Liang H, Mittler R (2003) The water-water cycle is essential for chloroplast protection in the absence of stress. *Journal of Biological Chemistry* 278: 38921-38925.
131. Steiner S, Schroter Y, Pfalz J, Pfannschmidt T (2011) Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiology* 157: 1043-1055.
132. Foyer CH, Halliwell B (1976) Presence of Glutathione and Glutathione Reductase in Chloroplasts - Proposed Role in Ascorbic-Acid Metabolism. *Planta* 133: 21-25.

133. Miyake C, Asada K (1992) Thylakoid-Bound Ascorbate Peroxidase in Spinach-Chloroplasts and Photoreduction of Its Primary Oxidation-Product Monodehydroascorbate Radicals in Thylakoids. *Plant and Cell Physiology* 33: 541-553.
134. Maruta T, Tanouchi A, Tamoi M, Yabuta Y, Yoshimura K, et al. (2010) Arabidopsis chloroplastic ascorbate peroxidase isoenzymes play a dual role in photoprotection and gene regulation under photooxidative stress. *Plant and Cell Physiology* 51: 190-200.
135. Rodriguez Milla MA, Maurer A, Rodriguez Huete A, Gustafson JP (2003) Glutathione peroxidase genes in Arabidopsis are ubiquitous and regulated by abiotic stresses through diverse signaling pathways. *Plant Journal* 36: 602-615.
136. Chang CC, Slesak I, Jorda L, Sotnikov A, Melzer M, et al. (2009) Arabidopsis chloroplastic glutathione peroxidases play a role in cross talk between photooxidative stress and immune responses. *Plant Physiology* 150: 670-683.
137. Zhai CZ, Zhao L, Yin LJ, Chen M, Wang QY, et al. (2013) Two wheat glutathione peroxidase genes whose products are located in chloroplasts improve salt and H₂O₂ tolerances in Arabidopsis. *PLoS One* 8: e73989.
138. Obara K, Sumi K, Fukuda H (2002) The use of multiple transcription starts causes the dual targeting of Arabidopsis putative monodehydroascorbate reductase to both mitochondria and chloroplasts. *Plant and Cell Physiology* 43: 697-705.
139. Creissen G, Reynolds H, Xue Y, Mullineaux P (1995) Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. *Plant Journal* 8: 167-175.
140. Chew O, Whelan J, Millar AH (2003) Molecular definition of the ascorbate-glutathione cycle in Arabidopsis mitochondria reveals dual targeting of antioxidant defenses in plants. *Journal of Biological Chemistry* 278: 46869-46877.
141. Yoshimura K, Yabuta Y, Ishikawa T, Shigeoka S (2002) Identification of a cis element for tissue-specific alternative splicing of chloroplast ascorbate peroxidase pre-mRNA in higher plants. *Journal of Biological Chemistry* 277: 40623-40632.
142. Kangasjarvi S, Lepisto A, Hannikainen K, Piippo M, Luomala EM, et al. (2008) Diverse roles for chloroplast stromal and thylakoid-bound ascorbate peroxidases in plant stress responses. *Biochemical Journal* 412: 275-285.
143. Dietz KJ, Jacob S, Oelze ML, Laxa M, Tognetti V, et al. (2006) The function of peroxiredoxins in plant organelle redox metabolism. *Journal of Experimental Botany* 57: 1697-1709.
144. Konig J, Baier M, Horling F, Kahmann U, Harris G, et al. (2002) The plant-specific function of 2-Cys peroxiredoxin-mediated detoxification of peroxides in the redox-hierarchy of photosynthetic electron flux. *Proceedings of the National Academy of Sciences of the United States of America* 99: 5738-5743.
145. Lamkemeyer P, Laxa M, Collin V, Li W, Finkemeier I, et al. (2006) Peroxiredoxin Q of Arabidopsis thaliana is attached to the thylakoids and functions in context of photosynthesis. *Plant Journal* 45: 968-981.
146. Serrato AJ, Perez-Ruiz JM, Spinola MC, Cejudo FJ (2004) A novel NADPH thioredoxin reductase, localized in the chloroplast, which deficiency causes hypersensitivity to abiotic stress in Arabidopsis thaliana. *Journal of Biological Chemistry* 279: 43821-43827.
147. Pulido P, Spinola MC, Kirchsteiger K, Guinea M, Pascual MB, et al. (2010) Functional analysis of the pathways for 2-Cys peroxiredoxin reduction in Arabidopsis thaliana chloroplasts. *Journal of Experimental Botany* 61: 4043-4054.

148. Puerto-Galan L, Perez-Ruiz JM, Guinea M, Cejudo FJ (2015) The contribution of NADPH thioredoxin reductase C (NTRC) and sulfiredoxin to 2-Cys peroxiredoxin overoxidation in *Arabidopsis thaliana* chloroplasts. *Journal of Experimental Botany*.
149. Awad J, Stotz HU, Fekete A, Krischke M, Engert C, et al. (2015) 2-Cysteine Peroxiredoxins and Thylakoid Ascorbate Peroxidase Create a Water-Water Cycle That Is Essential to Protect the Photosynthetic Apparatus under High Light Stress Conditions. *Plant Physiology* 167: 1592-1603.
150. Cooke MS, Evans MD, Dizdaroglu M, Lunec J (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB Journal* 17: 1195-1214.
151. Aller P, Rould MA, Hogg M, Wallace SS, Doublié S (2007) A structural rationale for stalling of a replicative DNA polymerase at the most common oxidative thymine lesion, thymine glycol. *Proceedings of the National Academy of Sciences of the United States of America* 104: 814-818.
152. Zeman MK, Cimprich KA (2014) Causes and consequences of replication stress. *Nature Cell Biology* 16: 2-9.
153. Shibutani S, Takeshita M, Grollman AP (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 349: 431-434.
154. Duncan BK, Miller JH (1980) Mutagenic deamination of cytosine residues in DNA. *Nature* 287: 560-561.
155. Degtyareva NP, Heyburn L, Sterling J, Resnick MA, Gordenin DA, et al. (2013) Oxidative stress-induced mutagenesis in single-strand DNA occurs primarily at cytosines and is DNA polymerase zeta-dependent only for adenines and guanines. *Nucleic Acids Research* 41: 8995-9005.
156. Zharkov DO (2008) Base excision DNA repair. *Cellular and Molecular Life Sciences* 65: 1544-1565.
157. Roldan-Arjona T, Garcia-Ortiz MV, Ruiz-Rubio M, Ariza RR (2000) cDNA cloning, expression and functional characterization of an *Arabidopsis thaliana* homologue of the *Escherichia coli* DNA repair enzyme endonuclease III. *Plant Molecular Biology* 44: 43-52.
158. Gutman BL, Niyogi KK (2009) Evidence for base excision repair of oxidative DNA damage in chloroplasts of *Arabidopsis thaliana*. *Journal of Biological Chemistry* 284: 17006-17012.
159. Ohtsubo T, Matsuda O, Iba K, Terashima I, Sekiguchi M, et al. (1998) Molecular cloning of AtMMH, an *Arabidopsis thaliana* ortholog of the *Escherichia coli* mutM gene, and analysis of functional domains of its product. *Molecular and General Genetics* 259: 577-590.
160. Murphy TM, Gao MJ (2001) Multiple forms of formamidopyrimidine-DNA glycosylase produced by alternative splicing in *Arabidopsis thaliana*. *Journal of Photochemistry and Photobiology B, Biology* 61: 87-93.
161. Bensen RJ, Warner HR (1987) Partial Purification and Characterization of Uracil-DNA Glycosylase Activity from Chloroplasts of *Zea mays* Seedlings. *Plant Physiology* 84: 1102-1106.
162. Babiychuk E, Kushnir S, Van Montagu M, Inze D (1994) The *Arabidopsis thaliana* apurinic endonuclease Arp reduces human transcription factors Fos and Jun. *Proceedings of the National Academy of Sciences of the United States of America* 91: 3299-3303.
163. Mitchell DL, Adair GM, Nairn RS (1989) Inhibition of transient gene expression in Chinese hamster ovary cells by triplet-sensitized UV-B irradiation of transfected DNA. *Photochemistry and Photobiology* 50: 639-646.
164. Villani G, Boiteux S, Radman M (1978) Mechanism of ultraviolet-induced mutagenesis: extent and fidelity of in vitro DNA synthesis on irradiated templates. *Proceedings of the National Academy of Sciences of the United States of America* 75: 3037-3041.

165. Britt AB, Chen JJ, Wykoff D, Mitchell D (1993) A UV-sensitive mutant of *Arabidopsis* defective in the repair of pyrimidine-pyrimidinone(6-4) dimers. *Science* 261: 1571-1574.
166. Boesch P, Weber-Lotfi F, Ibrahim N, Tarasenko V, Cosset A, et al. (2011) DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. *Biochimica et Biophysica Acta* 1813: 186-200.
167. Small GD, Greimann CS (1977) Photoreactivation and dark repair of ultraviolet light-induced pyrimidine dimers in chloroplast DNA. *Nucleic Acids Res* 4: 2893-2902.
168. Petersen JL, Lang DW, Small GD (1999) Cloning and characterization of a class II DNA photolyase from *Chlamydomonas*. *Plant Molecular Biology* 40: 1063-1071.
169. Petersen JL, Small GD (2001) A gene required for the novel activation of a class II DNA photolyase in *Chlamydomonas*. *Nucleic Acids Research* 29: 4472-4481.
170. Cannon GC, Hedrick LA, Heinhorst S (1995) Repair mechanisms of UV-induced DNA damage in soybean chloroplasts. *Plant Molecular Biology* 29: 1267-1277.
171. Stapleton AE, Thornber CS, Walbot V (1997) UV-B component of sunlight causes measurable damage in field-grown maize (*Zea mays* L): Developmental and cellular heterogeneity of damage and repair. *Plant Cell and Environment* 20: 279-290.
172. Takahashi M, Teranishi M, Ishida H, Kawasaki J, Takeuchi A, et al. (2011) Cyclobutane pyrimidine dimer (CPD) photolyase repairs ultraviolet-B-induced CPDs in rice chloroplast and mitochondrial DNA. *Plant Journal* 66: 433-442.
173. Chen JJ, Jiang CZ, Britt AB (1996) Little or No Repair of Cyclobutyl Pyrimidine Dimers Is Observed in the Organellar Genomes of the Young *Arabidopsis* Seedling. *Plant Physiology* 111: 19-25.
174. Kaiser G, Kleiner O, Beisswenger C, Batschauer A (2009) Increased DNA repair in *Arabidopsis* plants overexpressing CPD photolyase. *Planta* 230: 505-515.
175. Ahmad M, Jarillo JA, Klimczak LJ, Landry LG, Peng T, et al. (1997) An enzyme similar to animal type II photolyases mediates photoreactivation in *Arabidopsis*. *Plant Cell* 9: 199-207.
176. Kleine T, Lockhart P, Batschauer A (2003) An *Arabidopsis* protein closely related to *Synechocystis* cryptochrome is targeted to organelles. *Plant Journal* 35: 93-103.
177. Pokorny R, Klar T, Hennecke U, Carell T, Batschauer A, et al. (2008) Recognition and repair of UV lesions in loop structures of duplex DNA by DASH-type cryptochrome. *Proceedings of the National Academy of Sciences of the United States of America* 105: 21023-21027.
178. Lin C (2002) Blue light receptors and signal transduction. *Plant Cell* 14 Suppl: S207-225.
179. Lin C, Robertson DE, Ahmad M, Raibekas AA, Jorns MS, et al. (1995) Association of flavin adenine dinucleotide with the *Arabidopsis* blue light receptor CRY1. *Science* 269: 968-970.
180. Brautigam CA, Smith BS, Ma Z, Palnitkar M, Tomchick DR, et al. (2004) Structure of the photolyase-like domain of cryptochrome 1 from *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12142-12147.
181. Ahmad M, Cashmore AR (1993) HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* 366: 162-166.
182. Cashmore AR, Jarillo JA, Wu YJ, Liu D (1999) Cryptochromes: blue light receptors for plants and animals. *Science* 284: 760-765.
183. Brudler R, Hitomi K, Daiyasu H, Toh H, Kucho K, et al. (2003) Identification of a new cryptochrome class. Structure, function, and evolution. *Molecular Cell* 11: 59-67.
184. Huang Y, Baxter R, Smith BS, Partch CL, Colbert CL, et al. (2006) Crystal structure of cryptochrome 3 from *Arabidopsis thaliana* and its implications for photolyase activity. *Proceedings of the National Academy of Sciences of the United States of America* 103: 17701-17706.

185. Selby CP, Sancar A (2006) A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. *Proceedings of the National Academy of Sciences of the United States of America* 103: 17696-17700.
186. Deriano L, Roth DB (2013) Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annual Review of Genetics* 47: 433-455.
187. Helleday T, Lo J, van Gent DC, Engelward BP (2007) DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amst)* 6: 923-935.
188. McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in Genetics* 24: 529-538.
189. Durrenberger F, Rochaix JD (1991) Chloroplast ribosomal intron of *Chlamydomonas reinhardtii*: in vitro self-splicing, DNA endonuclease activity and in vivo mobility. *EMBO Journal* 10: 3495-3501.
190. Durrenberger F, Thompson AJ, Herrin DL, Rochaix JD (1996) Double strand break-induced recombination in *Chlamydomonas reinhardtii* chloroplasts. *Nucleic Acids Res* 24: 3323-3331.
191. Odom OW, Baek KH, Dani RN, Herrin DL (2008) *Chlamydomonas* chloroplasts can use short dispersed repeats and multiple pathways to repair a double-strand break in the genome. *Plant Journal* 53: 842-853.
192. Pannunzio NR, Li S, Watanabe G, Lieber MR (2014) Non-homologous end joining often uses microhomology: implications for alternative end joining. *DNA Repair (Amst)* 17: 74-80.
193. Roth DB, Wilson JH (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Molecular and Cellular Biology* 6: 4295-4304.
194. Kohl S, Bock R (2009) Transposition of a bacterial insertion sequence in chloroplasts. *Plant Journal* 58: 423-436.
195. Nussenzweig A, Nussenzweig MC (2007) A backup DNA repair pathway moves to the forefront. *Cell* 131: 223-225.
196. Bennardo N, Cheng A, Huang N, Stark JM (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genetics* 4: e1000110.
197. Costes A, Lambert SA (2012) Homologous recombination as a replication fork escort: fork-protection and recovery. *Biomolecules* 3: 39-71.
198. Lydeard JR, Lipkin-Moore Z, Sheu YJ, Stillman B, Burgers PM, et al. (2010) Break-induced replication requires all essential DNA replication factors except those specific for pre-RC assembly. *Genes and Development* 24: 1133-1144.
199. Llorente B, Smith CE, Symington LS (2008) Break-induced replication: what is it and what is it for? *Cell Cycle* 7: 859-864.
200. Hastings PJ, Ira G, Lupski JR (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genetics* 5: e1000327.
201. Hastings PJ, Lupski JR, Rosenberg SM, Ira G (2009) Mechanisms of change in gene copy number. *Nature Reviews: Genetics* 10: 551-564.
202. Cerutti H, Osman M, Grandoni P, Jagendorf AT (1992) A homolog of *Escherichia coli* RecA protein in plastids of higher plants. *Proceedings of the National Academy of Sciences of the United States of America* 89: 8068-8072.
203. Khazi FR, Edmondson AC, Nielsen BL (2003) An *Arabidopsis* homologue of bacterial RecA that complements an *E. coli* recA deletion is targeted to plant mitochondria. *Molecular Genetics and Genomics* 269: 454-463.
204. Nakazato E, Fukuzawa H, Tabata S, Takahashi H, Tanaka K (2003) Identification and expression analysis of cDNA encoding a chloroplast recombination protein REC1, the chloroplast RecA

- homologue in *Chlamydomonas reinhardtii*. *Bioscience, Biotechnology, and Biochemistry* 67: 2608-2613.
205. Inouye T, Odahara M, Fujita T, Hasebe M, Sekine Y (2008) Expression and complementation analyses of a chloroplast-localized homolog of bacterial RecA in the moss *Physcomitrella patens*. *Bioscience, Biotechnology, and Biochemistry* 72: 1340-1347.
 206. Odahara M, Inouye T, Fujita T, Hasebe M, Sekine Y (2007) Involvement of mitochondrial-targeted RecA in the repair of mitochondrial DNA in the moss, *Physcomitrella patens*. *Genes and Genetic Systems* 82: 43-51.
 207. Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA (2007) Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* 19: 1251-1264.
 208. Cerutti H, Jagendorf AT (1993) DNA Strand-Transfer Activity in Pea (*Pisum sativum* L.) Chloroplasts. *Plant Physiology* 102: 145-153.
 209. Cerutti H, Johnson AM, Boynton JE, Gillham NW (1995) Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of *Escherichia coli* RecA. *Molecular and Cellular Biology* 15: 3003-3011.
 210. Rowan BA, Oldenburg DJ, Bendich AJ (2010) RecA maintains the integrity of chloroplast DNA molecules in *Arabidopsis*. *Journal of Experimental Botany* 61: 2575-2588.
 211. Maréchal A, Brisson N (2010) Recombination and the maintenance of plant organelle genome stability. *New Phytologist* 186: 299-317.
 212. Odahara M, Kuroiwa H, Kuroiwa T, Sekine Y (2009) Suppression of repeat-mediated gross mitochondrial genome rearrangements by RecA in the moss *Physcomitrella patens*. *Plant Cell* 21: 1182-1194.
 213. Miller-Messmer M, Kuhn K, Bichara M, Le Ret M, Imbault P, et al. (2012) RecA-dependent DNA repair results in increased heteroplasmy of the *Arabidopsis* mitochondrial genome. *Plant Physiology* 159: 211-226.
 214. Mukaihara T, Enomoto M (1997) Deletion formation between the two *Salmonella typhimurium* flagellin genes encoded on the mini F plasmid: *Escherichia coli* ssb alleles enhance deletion rates and change hot-spot preference for deletion endpoints. *Genetics* 145: 563-572.
 215. Reddy M, Gowrishankar J (1997) Identification and characterization of ssb and uup mutants with increased frequency of precise excision of transposon Tn10 derivatives: nucleotide sequence of uup in *Escherichia coli*. *Journal of Bacteriology* 179: 2892-2899.
 216. Shereda RD, Kozlov AG, Lohman TM, Cox MM, Keck JL (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology* 43: 289-318.
 217. Krause K, Kilbiński I, Mulisch M, Rodiger A, Schafer A, et al. (2005) DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Letters* 579: 3707-3712.
 218. Cappadocia L, Parent JS, Zampini E, Lepage E, Sygusch J, et al. (2012) A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage. *Nucleic Acids Res* 40: 258-269.
 219. Abdelnoor RV, Yule R, Elo A, Christensen AC, Meyer-Gauen G, et al. (2003) Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. *Proceedings of the National Academy of Sciences of the United States of America* 100: 5968-5973.

220. Xu YZ, Arrieta-Montiel MP, Viridi KS, de Paula WB, Widhalm JR, et al. (2011) MutS HOMOLOG1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light. *Plant Cell* 23: 3428-3441.
221. Arrieta-Montiel MP, Shedge V, Davila J, Christensen AC, Mackenzie SA (2009) Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* 183: 1261-1268.
222. Rédei GP, Plurad SB (1973) Hereditary Structural Alterations of Plastids Induced by a Nuclear Mutator Gene in Arabidopsis. *Protoplasma* 77: 361-380.
223. Zaegel V, Guermann B, Le Ret M, Andres C, Meyer D, et al. (2006) The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in Arabidopsis. *Plant Cell* 18: 3548-3563.
224. Odahara M, Masuda Y, Sato M, Wakazaki M, Harada C, et al. (2015) RECG maintains plastid and mitochondrial genome stability by suppressing extensive recombination between short dispersed repeats. *PLoS Genetics* 11: e1005080.
225. Yeeles JT, Poli J, Mariani KJ, Pasero P (2013) Rescuing stalled or damaged replication forks. *Cold Spring Harb Perspect Biol* 5: a012815.
226. Samach A, Melamed-Bessudo C, Avivi-Ragolski N, Pietrokovski S, Levy AA (2011) Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. *Plant Cell* 23: 4266-4279.
227. Janicka S, Kuhn K, Le Ret M, Bonnard G, Imbault P, et al. (2012) A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *Plant Journal*.
228. Tuppen HA, Blakely EL, Turnbull DM, Taylor RW (2010) Mitochondrial DNA mutations and human disease. *Biochimica et Biophysica Acta* 1797: 113-128.
229. Baena-Gonzalez E, Allahverdiyeva Y, Svab Z, Maliga P, Josse EM, et al. (2003) Deletion of the tobacco plastid psbA gene triggers an upregulation of the thylakoid-associated NAD(P)H dehydrogenase complex and the plastid terminal oxidase (PTOX). *Plant Journal* 35: 704-716.
230. Pfannschmidt T, Nilsson A, Allen JF (1999) Photosynthetic control of chloroplast gene expression. *Nature* 397: 625-628.
231. Huner NP, Bode R, Dahal K, Hollis L, Rosso D, et al. (2012) Chloroplast redox imbalance governs phenotypic plasticity: the "grand design of photosynthesis" revisited. *Front Plant Sci* 3: 255.
232. Przybyla D, Gobel C, Imboden A, Hamberg M, Feussner I, et al. (2008) Enzymatic, but not non-enzymatic, 1O₂-mediated peroxidation of polyunsaturated fatty acids forms part of the EXECUTER1-dependent stress response program in the flu mutant of Arabidopsis thaliana. *Plant Journal* 54: 236-248.
233. Triantaphylides C, Krischke M, Hoeberichts FA, Ksas B, Gresser G, et al. (2008) Singlet oxygen is the major reactive oxygen species involved in photooxidative damage to plants. *Plant Physiology* 148: 960-968.
234. Mittler R (2002) Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* 7: 405-410.
235. Miller G, Shulaev V, Mittler R (2008) Reactive oxygen signaling and abiotic stress. *Physiologia Plantarum* 133: 481-489.
236. Torres MA (2010) ROS in biotic interactions. *Physiologia Plantarum* 138: 414-429.
237. Ramel F, Ksas B, Akkari E, Mialoundama AS, Monnet F, et al. (2013) Light-Induced Acclimation of the Arabidopsis chlorina1 Mutant to Singlet Oxygen. *Plant Cell* 25: 1445-1462.
238. Zaltsman A, Ori N, Adam Z (2005) Two types of FtsH protease subunits are required for chloroplast biogenesis and Photosystem II repair in Arabidopsis. *Plant Cell* 17: 2782-2790.

239. Stroher E, Dietz KJ (2008) The dynamic thiol-disulphide redox proteome of the *Arabidopsis thaliana* chloroplast as revealed by differential electrophoretic mobility. *Physiologia Plantarum* 133: 566-583.
240. Sun X, Fu T, Chen N, Guo J, Ma J, et al. (2010) The stromal chloroplast Deg7 protease participates in the repair of photosystem II after photoinhibition in *Arabidopsis*. *Plant Physiology* 152: 1263-1273.
241. Arsova B, Hoja U, Wimmelbacher M, Greiner E, Ustun S, et al. (2010) Plastidial thioredoxin z interacts with two fructokinase-like proteins in a thiol-dependent manner: evidence for an essential role in chloroplast development in *Arabidopsis* and *Nicotiana benthamiana*. *Plant Cell* 22: 1498-1515.
242. Kindgren P, Kremnev D, Blanco NE, de Dios Barajas Lopez J, Fernandez AP, et al. (2012) The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. *Plant Journal* 70: 279-291.
243. Drlica K, Zhao X (1997) DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiology and Molecular Biology Reviews* 61: 377-392.
244. Cupp JD, Nielsen BL (2012) *Arabidopsis thaliana* organellar DNA polymerase IB mutants exhibit reduced mtDNA levels with a decrease in mitochondrial area density. *Physiologia Plantarum* 149: 91-103.
245. Rédei GP (1963) Somatic Instability Caused by a Cysteine-Sensitive Gene in *Arabidopsis*. *Science* 139: 767-769.
246. Martínez-Zapater JM (1993) Genetic-Analysis of Variegated Mutants in *Arabidopsis*. *Journal of Heredity* 84: 138-140.
247. Hardy CD, Cozzarelli NR (2003) Alteration of *Escherichia coli* topoisomerase IV to novobiocin resistance. *Antimicrobial Agents and Chemotherapy* 47: 941-947.
248. Drier SM, Fryer MJ, Mullineaux PM, Baker NR (2009) Imaging of reactive oxygen species in vivo. *Methods in Molecular Biology* 479: 109-116.
249. Weigel D, Glazebrook J (2002) *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Lab Press, Cold Spring Harbor, NY.
250. Yu F, Fu A, Aluru M, Park S, Xu Y, et al. (2007) Variegation mutants and mechanisms of chloroplast biogenesis. *Plant, Cell & Environment* 30: 350-365.
251. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant Journal* 37: 914-939.
252. Ecker JR (1995) The ethylene signal transduction pathway in plants. *Science* 268: 667-675.
253. Cheng MC, Liao PM, Kuo WW, Lin TP (2013) The *Arabidopsis* ETHYLENE-RESPONSE-FACTOR1 Regulates Abiotic-Stress-Responsive Gene Expression by Binding to Different cis-Acting Elements in Response to Different Stress Signals. *Plant Physiology*.
254. Usadel B, Nagel A, Steinhauser D, Gibon Y, Blasing OE, et al. (2006) PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* 7: 535.
255. Chia LS, McRae DG, Thompson JE (1982) Light-Dependence of Paraquat-Initiated Membrane Deterioration in Bean-Plants - Evidence for the Involvement of Superoxide. *Physiologia Plantarum* 56: 492-499.
256. Ramel F, Sulmon C, Bogard M, Couée I, Gouesbet G (2009) Differential patterns of reactive oxygen species and antioxidative mechanisms during atrazine injury and sucrose-induced tolerance in *Arabidopsis thaliana* plantlets. *BMC Plant Biology* 9: 28.

257. Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, et al. (1998) Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant Journal* 16: 433-442.
258. Sakamoto H, Maruyama K, Sakuma Y, Meshi T, Iwabuchi M, et al. (2004) Arabidopsis Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiology* 136: 2734-2746.
259. Walley JW, Kelley DR, Nestorova G, Hirschberg DL, Dehesh K (2010) Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiology* 152: 866-875.
260. Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF (2005) Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of Arabidopsis. *Plant Journal* 41: 195-211.
261. Gonzalez-Perez S, Gutierrez J, Garcia-Garcia F, Osuna D, Dopazo J, et al. (2011) Early transcriptional defense responses in Arabidopsis cell suspension culture under high-light conditions. *Plant Physiol* 156: 1439-1456.
262. Aluru MR, Zola J, Foudree A, Rodermeier SR (2009) Chloroplast photooxidation-induced transcriptome reprogramming in Arabidopsis *immutans* white leaf sectors. *Plant Physiology* 150: 904-923.
263. Miura E, Kato Y, Sakamoto W (2010) Comparative transcriptome analysis of green/white variegated sectors in Arabidopsis yellow variegated2: responses to oxidative and other stresses in white sectors. *Journal of Experimental Botany* 61: 2433-2445.
264. Ogawa T, Ueda Y, Yoshimura K, Shigeoka S (2005) Comprehensive analysis of cytosolic Nudix hydrolases in Arabidopsis thaliana. *Journal of Biological Chemistry* 280: 25277-25283.
265. Sweetlove LJ, Heazlewood JL, Herald V, Holtzapffel R, Day DA, et al. (2002) The impact of oxidative stress on Arabidopsis mitochondria. *Plant Journal* 32: 891-904.
266. Arteca JM, Arteca RN (1999) A multi-responsive gene encoding 1-aminocyclopropane-1-carboxylate synthase (ACS6) in mature Arabidopsis leaves. *Plant Molecular Biology* 39: 209-219.
267. Kimura M, Yoshizumi T, Manabe K, Yamamoto YY, Matsui M (2001) Arabidopsis transcriptional regulation by light stress via hydrogen peroxide-dependent and -independent pathways. *Genes to Cells* 6: 607-617.
268. Petersen K, Schöttler MA, Karcher D, Thiele W, Bock R (2011) Elimination of a group II intron from a plastid gene causes a mutant phenotype. *Nucleic Acids Research* 39: 5181-5192.
269. Asada K (1999) THE WATER-WATER CYCLE IN CHLOROPLASTS: Scavenging of Active Oxygens and Dissipation of Excess Photons. *Annual Review of Plant Physiology and Plant Molecular Biology* 50: 601-639.
270. Ishikawa K, Takenaga K, Akimoto M, Koshikawa N, Yamaguchi A, et al. (2008) ROS-generating mitochondrial DNA mutations can regulate tumor cell metastasis. *Science* 320: 661-664.
271. Fan W, Lin CS, Potluri P, Procaccio V, Wallace DC (2012) mtDNA lineage analysis of mouse L-cell lines reveals the accumulation of multiple mtDNA mutants and intermolecular recombination. *Genes and Development* 26: 384-394.
272. Avni A, Edelman M, Rachailovich I, Aviv D, Fluhr R (1989) A Point Mutation in the Gene for the Large Subunit of Ribulose 1,5-Bisphosphate Carboxylase Oxygenase Affects Holoenzyme Assembly in Nicotiana-Tabacum. *EMBO Journal* 8: 1915-1918.
273. Triboush SO, Danilenko NG, Ulitcheva II, Davydenko OG (1999) Location of induced mutations and reversions in the chloroplast genome of Helianthus annuus. *Plant Growth Regulation* 27: 75-81.

274. Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology* 55: 373-399.
275. Mittler R, Vanderauwera S, Gollery M, Van Breusegem F (2004) Reactive oxygen gene network of plants. *Trends in Plant Science* 9: 490-498.
276. Rosso D, Bode R, Li W, Krol M, Saccon D, et al. (2009) Photosynthetic redox imbalance governs leaf sectoring in the *Arabidopsis thaliana* variegation mutants *immutans*, *spotty*, *var1*, and *var2*. *Plant Cell* 21: 3473-3492.
277. Rentel MC, Knight MR (2004) Oxidative stress-induced calcium signaling in *Arabidopsis*. *Plant Physiology* 135: 1471-1479.
278. Shi Y, Tian S, Hou L, Huang X, Zhang X, et al. (2012) Ethylene signaling negatively regulates freezing tolerance by repressing expression of CBF and type-A ARR genes in *Arabidopsis*. *Plant Cell* 24: 2578-2595.
279. Galon Y, Finkler A, Fromm H (2010) Calcium-regulated transcription in plants. *Mol Plant* 3: 653-669.
280. Prasad TK, Anderson MD, Martin BA, Stewart CR (1994) Evidence for Chilling-Induced Oxidative Stress in Maize Seedlings and a Regulatory Role for Hydrogen Peroxide. *Plant Cell* 6: 65-74.
281. Coll NS, Danon A, Meurer J, Cho WK, Apel K (2009) Characterization of *soldat8*, a suppressor of singlet oxygen-induced cell death in *Arabidopsis* seedlings. *Plant and Cell Physiology* 50: 707-718.
282. Šimková K, Kim C, Gacek K, Baruah A, Laloi C, et al. (2012) The chloroplast division mutant *caa33* of *Arabidopsis thaliana* reveals the crucial impact of chloroplast homeostasis on stress acclimation and retrograde plastid-to-nucleus signaling. *Plant Journal* 69: 701-712.
283. Gordon MJ, Carmody M, Albrecht V, Pogson B (2013) Systemic and Local Responses to Repeated HL Stress-Induced Retrograde Signaling in *Arabidopsis*. *Front Plant Sci* 3: 303.
284. Porra RJ, Thompson WA, Kriedemann PE (1989) Determination of Accurate Extinction Coefficients and Simultaneous-Equations for Assaying Chlorophyll-a and Chlorophyll-B Extracted with 4 Different Solvents - Verification of the Concentration of Chlorophyll Standards by Atomic-Absorption Spectroscopy. *Biochimica et Biophysica Acta* 975: 384-394.
285. Leshem Y, Melamed-Book N, Cagnac O, Ronen G, Nishri Y, et al. (2006) Suppression of *Arabidopsis* vesicle-SNARE expression inhibited fusion of H₂O₂-containing vesicles with tonoplast and increased salt tolerance. *Proceedings of the National Academy of Sciences of the United States of America* 103: 18008-18013.
286. Cottage A, Mott EK, Kempster JA, Gray JC (2010) The *Arabidopsis* plastid-signalling mutant *gun1* (*genomes uncoupled1*) shows altered sensitivity to sucrose and abscisic acid and alterations in early seedling development. *Journal of Experimental Botany* 61: 3773-3786.
287. Goecks J, Nekrutenko A, Taylor J, Galaxy T (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11: R86.
288. Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, et al. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* 2008: 420747.
289. Lepage E, Zampini E, Brisson N (2013) Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in *Arabidopsis*. *Plant Physiology* 163: 867-881.
290. Deng SK, Gibb B, de Almeida MJ, Greene EC, Symington LS (2014) RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nature Structural & Molecular Biology* 21: 405-412.

291. Seigneur M, Ehrlich SD, Michel B (2000) RuvABC-dependent double-strand breaks in dnaBts mutants require recA. *Molecular Microbiology* 38: 565-574.
292. Robu ME, Inman RB, Cox MM (2001) RecA protein promotes the regression of stalled replication forks in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 98: 8211-8218.
293. Vermulst M, Wanagat J, Kujoth GC, Bielas JH, Rabinovitch PS, et al. (2008) DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. *Nature Genetics* 40: 392-394.
294. Ameer A, Stewart JB, Freyer C, Hagstrom E, Ingman M, et al. (2011) Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genetics* 7: e1002028.
295. Lundin C, Erixon K, Arnaudeau C, Schultz N, Jenssen D, et al. (2002) Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. *Molecular and Cellular Biology* 22: 5869-5878.
296. Lee JA, Carvalho CM, Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 131: 1235-1247.
297. Mizuno K, Lambert S, Baldacci G, Murray JM, Carr AM (2009) Nearby inverted repeats fuse to generate acentric and dicentric palindromic chromosomes by a replication template exchange mechanism. *Genes and Development* 23: 2876-2886.
298. Paek AL, Kaochar S, Jones H, Elezaby A, Shanks L, et al. (2009) Fusion of nearby inverted repeats by a replication-based mechanism leads to formation of dicentric and acentric chromosomes that cause genome instability in budding yeast. *Genes and Development* 23: 2861-2875.
299. Mizuno K, Miyabe I, Schalbetter SA, Carr AM, Murray JM (2013) Recombination-restarted replication makes inverted chromosome fusions at inverted repeats. *Nature* 493: 246-249.
300. Kugelberg E, Kofoed E, Andersson DI, Lu Y, Mellor J, et al. (2010) The tandem inversion duplication in *Salmonella enterica*: selection drives unstable precursors to final mutation types. *Genetics* 185: 65-80.
301. Seier T, Zilberberg G, Zeiger DM, Lovett ST (2012) Azidothymidine and other chain terminators are mutagenic for template-switch-generated genetic mutations. *Proceedings of the National Academy of Sciences of the United States of America* 109: 6171-6174.
302. Weigel D, Mott R (2009) The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biol* 10: 107.
303. Slager J, Kjos M, Attaiech L, Veening JW (2014) Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell* 157: 395-406.
304. Wu J, Liu B, Cheng F, Ramchiary N, Choi SR, et al. (2012) Sequencing of chloroplast genome using whole cellular DNA and solexa sequencing technology. *Front Plant Sci* 3: 243.
305. Ferrarini M, Moretto M, Ward JA, Surbanovski N, Stevanovic V, et al. (2013) An evaluation of the PacBio RS platform for sequencing and de novo assembly of a chloroplast genome. *BMC Genomics* 14: 670.
306. Small ID, Isaac PG, Leaver CJ (1987) Stoichiometric differences in DNA molecules containing the *atpA* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. *EMBO Journal* 6: 865-869.
307. Ogihara Y, Terachi T, Sasakuma T (1988) Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proceedings of the National Academy of Sciences of the United States of America* 85: 8573-8577.
308. Bender A, Krishnan KJ, Morris CM, Taylor GA, Reeve AK, et al. (2006) High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nature Genetics* 38: 515-517.

309. Kraytsberg Y, Kudryavtseva E, McKee AC, Geula C, Kowall NW, et al. (2006) Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nature Genetics* 38: 518-520.
310. Moslemi AR, Lindberg C, Oldfors A (1997) Analysis of multiple mitochondrial DNA deletions in inclusion body myositis. *Human Mutation* 10: 381-386.
311. Ju YS, Alexandrov LB, Gerstung M, Martincorena I, Nik-Zainal S, et al. (2014) Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife* 3.
312. Pui CH, Carroll AJ, Raimondi SC, Schell MJ, Head DR, et al. (1992) Isochromosomes in childhood acute lymphoblastic leukemia: a collaborative study of 83 cases. *Blood* 79: 2384-2391.
313. Hermetz KE, Newman S, Conneely KN, Martin CL, Ballif BC, et al. (2014) Large inverted duplications in the human genome form via a fold-back mechanism. *PLoS Genetics* 10: e1004139.
314. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653-657.
315. Whitlock MC (2005) Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology* 18: 1368-1373.
316. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, et al. (2010) Manipulation of FASTQ data with Galaxy. *Bioinformatics* 26: 1783-1785.
317. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-1760.
318. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
319. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.
320. Little SE (1997) Large- or Small-Scale Isolation of Chloroplasts Using the Avanti® J Series of High Performance Centrifuges. Beckman Instruments, Inc.
321. Bilski P, Belanger AG, Chignell CF (2002) Photosensitized oxidation of 2',7'-dichlorofluorescein: singlet oxygen does not contribute to the formation of fluorescent oxidation product 2',7'-dichlorofluorescein. *Free Radical Biology and Medicine* 33: 938-946.
322. Eruslanov E, Kusmartsev S (2010) Identification of ROS using oxidized DCFDA and flow-cytometry. *Methods in Molecular Biology* 594: 57-72.
323. LeBel CP, Ischiropoulos H, Bondy SC (1992) Evaluation of the probe 2',7'-dichlorofluorescein as an indicator of reactive oxygen species formation and oxidative stress. *Chemical Research in Toxicology* 5: 227-231.
324. Atkinson J, McGlynn P (2009) Replication fork reversal and the maintenance of genome stability. *Nucleic Acids Research* 37: 3475-3492.
325. Yeeles JT, Marians KJ (2011) The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science* 334: 235-238.
326. Neelsen KJ, Lopes M (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nature Reviews: Molecular Cell Biology* 16: 207-220.
327. Ahmed A, Podemski L (1998) Observations on template switching during DNA replication through long inverted repeats. *Gene* 223: 187-194.
328. Bi X, Liu LF (1996) DNA rearrangement mediated by inverted repeats. *Proceedings of the National Academy of Sciences of the United States of America* 93: 819-823.
329. Courcelle J, Donaldson JR, Chow KH, Courcelle CT (2003) DNA damage-induced replication fork regression and processing in *Escherichia coli*. *Science* 299: 1064-1067.

330. Cox MM (2007) Motoring along with the bacterial RecA protein. *Nature Reviews: Molecular Cell Biology* 8: 127-138.
331. Wang TF, Chen LT, Wang AH (2008) Right or left turn? RecA family protein filaments promote homologous recombination through clockwise axial rotation. *Bioessays* 30: 48-56.
332. Jeon H, Jin YM, Choi MH, Lee H, Kim M (2013) Chloroplast-targeted bacterial RecA proteins confer tolerance to chloroplast DNA damage by methyl viologen or UV-C radiation in tobacco (*Nicotiana tabacum*) plants. *Physiologia Plantarum* 147: 218-233.
333. Buss JA, Kimura Y, Bianco PR (2008) RecG interacts directly with SSB: implications for stalled replication fork regression. *Nucleic Acids Research* 36: 7029-7042.
334. Lecointe F, Serena C, Velten M, Costes A, McGovern S, et al. (2007) Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks. *EMBO Journal* 26: 4239-4251.
335. Lenhart JS, Brandes ER, Schroeder JW, Sorenson RJ, Showalter HD, et al. (2014) RecO and RecR are necessary for RecA loading in response to DNA damage and replication fork stress. *Journal of Bacteriology* 196: 2851-2860.
336. Cappadocia L, Parent JS, Sygusch J, Brisson N (2013) A family portrait: structural comparison of the Whirly proteins from *Arabidopsis thaliana* and *Solanum tuberosum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69: 1207-1211.
337. Paek AL, Jones H, Kaochar S, Weinert T (2010) The role of replication bypass pathways in dicentric chromosome formation in budding yeast. *Genetics* 186: 1161-1173.
338. Benson FE, West SC (1994) Substrate specificity of the *Escherichia coli* RuvC protein. Resolution of three- and four-stranded recombination intermediates. *Journal of Biological Chemistry* 269: 5195-5201.
339. Jensch F, Kemper B (1986) Endonuclease VII resolves Y-junctions in branched DNA in vitro. *EMBO Journal* 5: 181-189.
340. Bendich AJ (2010) The end of the circle for yeast mitochondrial DNA. *Molecular Cell* 39: 831-832.
341. Gerhold JM, Aun A, Sedman T, Joers P, Sedman J (2010) Strand invasion structures in the inverted repeat of *Candida albicans* mitochondrial DNA reveal a role for homologous recombination in replication. *Molecular Cell* 39: 851-861.
342. Pohjoismaki JL, Goffart S, Tynjismaa H, Willcox S, Ide T, et al. (2009) Human heart mitochondrial DNA is organized in complex catenated networks containing abundant four-way junctions and replication forks. *Journal of Biological Chemistry* 284: 21446-21457.
343. Pohjoismaki JL, Goffart S (2011) Of circles, forks and humanity: Topological organisation and replication of mammalian mitochondrial DNA. *Bioessays* 33: 290-299.
344. Hori A, Yoshida M, Shibata T, Ling F (2009) Reactive oxygen species regulate DNA copy number in isolated yeast mitochondria by triggering recombination-mediated replication. *Nucleic Acids Research* 37: 749-761.
345. Pohjoismaki JL, Williams SL, Boettger T, Goffart S, Kim J, et al. (2013) Overexpression of Twinkle-helicase protects cardiomyocytes from genotoxic stress caused by reactive oxygen species. *Proceedings of the National Academy of Sciences of the United States of America* 110: 19408-19413.
346. Pohjoismaki JL, Goffart S, Taylor RW, Turnbull DM, Suomalainen A, et al. (2010) Developmental and pathological changes in the human cardiac muscle mitochondrial DNA organization, replication and copy number. *PLoS One* 5: e10426.
347. Kato Y, Miura E, Matsushima R, Sakamoto W (2007) White leaf sectors in yellow variegated2 are formed by viable cells with undifferentiated plastids. *Plant Physiology* 144: 952-960.

348. Sakamoto W, Uno Y, Zhang Q, Miura E, Kato Y, et al. (2009) Arrested differentiation of proplastids into chloroplasts in variegated leaves characterized by plastid ultrastructure and nucleoid morphology. *Plant and Cell Physiology* 50: 2069-2083.
349. Wang S, Wen R, Shi X, Lambrecht A, Wang H, et al. (2011) RAD5a and REV3 function in two alternative pathways of DNA-damage tolerance in *Arabidopsis*. *DNA Repair (Amst)* 10: 620-628.
350. Olivero OA (2007) Mechanisms of genotoxicity of nucleoside reverse transcriptase inhibitors. *Environmental and Molecular Mutagenesis* 48: 215-223.
351. Cooper DL, Lovett ST (2011) Toxicity and tolerance mechanisms for azidothymidine, a replication gap-promoting agent, in *Escherichia coli*. *DNA Repair (Amst)* 10: 260-270.
352. McGlynn P, Lloyd RG (2001) Rescue of stalled replication forks by RecG: simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and Holliday junction formation. *Proceedings of the National Academy of Sciences of the United States of America* 98: 8227-8234.
353. McGlynn P, Lloyd RG, Marians KJ (2001) Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. *Proceedings of the National Academy of Sciences of the United States of America* 98: 8235-8240.
354. Manosas M, Perumal SK, Bianco PR, Ritort F, Benkovic SJ, et al. (2013) RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue. *Nat Commun* 4: 2368.
355. Rowan BA, Oldenburg DJ, Bendich AJ (2007) A high-throughput method for detection of DNA in chloroplasts using flow cytometry. *Plant Methods* 3: 5.
356. Sage JM, Knight KL (2013) Human Rad51 promotes mitochondrial DNA synthesis under conditions of increased replication stress. *Mitochondrion* 13: 350-356.
357. Ruhanen H, Borrie S, Szabadkai G, Tynismaa H, Jones AW, et al. (2010) Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochimica et Biophysica Acta* 1803: 931-939.